A tutorial on fiber Kerr nonlinearity effect and its compensation in optical communication systems

OPEN ACCESS
IOP Publishing
Journal of Optics
J. Opt. **23** (2021) 123502 (24pp)
https://doi.org/10.1088/2040-8986/ac362a

Tutorial

A tutorial on fiber Kerr nonlinearity effect and its compensation in optical communication systems

Sunish Kumar Orappanpara Soman

School of Engineering, Ulster University, Jordanstown, Northern Ireland BT37 0QB, United Kingdom

E-mail: S.Orappanpara_Soman@ulster.ac.uk

Received 24 July 2021, revised 18 October 2021
Accepted for publication 3 November 2021
Published 22 November 2021

Abstract

The advent of silica-based low-cost standard single-mode fibers revolutionized the whole communication industry. The deployment of optical fibers in the networks induces a paradigm shift in the communication technologies used for long-haul information transfer. However, the communication using the optical fibers is affected by several linear and nonlinear effects. The most common linear effects are attenuation and chromatic dispersion, whereas the dominant nonlinear effect is the Kerr effect. The Kerr effect induces a power-dependent nonlinear distortion for the signal propagating in the optical fiber. The detrimental effects of the Kerr nonlinearity limit the capacity of long-haul optical communication systems. Fiber Kerr nonlinearity compensation using digital signal processing (DSP) techniques has been well investigated over several years. In this paper, we provide a comprehensive tutorial, including the fundamental mathematical analysis, on the characteristics of the optical fiber channel, the origin of the Kerr nonlinearity effect, the theory of the pulse propagation in the optical fiber, and the numerical and analytical tools for solving the pulse propagation equation. In addition, we provide a concise review of various DSP techniques for fiber nonlinearity compensation, such as digital back-propagation, Volterra series-based nonlinearity equalization, perturbation theory-based nonlinearity compensation, and phase conjugation. We also carry out numerical simulation and the complexity evaluation of the selected nonlinearity compensation techniques.

Keywords: digital back-propagation, digital signal processing, Kerr nonlinearity, optical fiber communications, perturbation theory, phase-conjugation, Volterra equalizer

(Some figures may appear in colour only in the online journal)

1. Introduction

The modern high capacity core communication network uses optical fibers to transmit information from one point to another as modulated light pulses. The optical fiber is a dielectric cylindrical waveguide made up of low-cost material silica [1]. The basic phenomenon responsible for guiding the light pulses inside the optical fiber is total internal reflection [1]. The uncladded optical fibers manufactured in the early 1920s were not suitable for information transfer at long distances [1]. In the 1950s, the use of the cladding layer was a starting point for the field of fiber optics and led to a significant improvement in the fiber characteristics [1]. During the 1960s, the transmission of images through the glass fibers demonstrated the significant

development of the field of fiber optics [1, 2]. The fibers manufactured during that time had a drawback of extreme power loss of $\ggg$ 1000 dB km^{-1} while the optical light is transmitted. Further research efforts in the 1970s reduced the loss of the silica-based fibers below 20 dB km^{-1} [1]. During the late 1970s, advancements in fiber fabrication technology significantly reduced the fiber loss down to 0.2 dB km^{-1} in the 1550 nm wavelength range [1]. That eventually led to a breakthrough in the field of optical fiber communication systems [1]. During the 1990s, the optical fibers were doped with rare-earth elements such as Erbium. That led to the development of the Erbium-doped fiber amplifiers (EDFAs) and fiber lasers [1, 3].

The invention of EDFAs and cost-effective manufacturing of standard single-mode fibers (SSMFs) with less than 0.2 dB km^{-1} attenuation marked the beginning of commercially viable fiber-optic communication systems. Since the deployment, the optical communication systems underwent several technical evolutions to fulfil the requirements of high-speed communications. The emergence of bandwidth-hungry applications, such as cloud services and virtual reality, has fueled the global network traffic increase to a large extent [4–7]. Furthermore, human-centered applications like video gaming and the exchange of multimedia content via smartphones are among the most bandwidth-consuming applications. That leads to a strong requirement for an increase in the access network capacity, and consequently, for the core optical network capacities to meet such ever-increasing traffic demands [7].

The deployment of wavelength division multiplexing (WDM) technology, which enables the multiplexing of several optical signals in the same fiber, dramatically increases the fiber capacity [8, 9]. A few years ago, an optical superchannel system has been introduced for 400 Gb s^{-1} or 1 Tb s^{-1} optical transmission systems [3–16]. This system effectively splits the given WDM channel into smaller subcarriers separated by smaller guard-bands and transmitted as a single entity along the optical network. The superchannel technique has several benefits over the single-carrier systems [16]. It has lower requirements in terms of optical signal-to-noise ratio (OSNR) and analog-to-digital converters/digital-to-analog converters bandwidth [16]. In comparison with single-carrier 400 Gb s^{-1} and 1 Tb s^{-1} transmission systems, the superchannel systems exhibit better transmission performance. However, the transmission impairments in the optical fiber degrade the performance of the high data-rate optical superchannel systems. For example, the chromatic dispersion (CD) introduces a frequency-dependent phase shift to the signals and acts as a primary limiting factor in provisioning a reliable long-haul optical communication link [17]. It is noteworthy that there are various optical and electrical techniques available to combat the adverse effect of the CD in a long-haul optical communication system [17–21].

Another significant impairment that limits the transmission performance of the long-haul optical communication system is the fiber nonlinearity [7]. The nonlinearity effects in the optical fiber are due to an electro-optic effect, referred to as the Kerr effect, which arises from the dependence of the optical fiber refractive index on the transmit signal power [2, 7]. That makes the optical fiber channel different from other transmission media used for the information transfer. In a linear transmission medium, the information signals are usually perturbed by the additive noise, which generally results in channel capacities monotonically increasing with transmit power, thereby a corresponding increase in the SNR. However, the detrimental effects of Kerr-induced signal nonlinear distortions grow at a faster rate than the SNR capacity gain at higher launch powers. That in turn leads the channel capacity to be a nonmonotonic function of the transmit launch power with a maximum value at a particular launch power termed as optimum launch power. The achievable transmission rate decreases rapidly beyond the optimal power point as the launch power increases due to the corresponding increase in the Kerr-induced signal nonlinear distortions [2].

The transmission performance of the single-channel optical communication systems is mainly limited by the intra-channel Kerr nonlinearity effect. The intra-channel nonlinear interactions can be categorized into three types: (a) self-phase modulation (SPM), (b) intra-channel cross-phase modulation (IXPM), and (c) intra-channel four-wave mixing (IFWM). It is important to mention that the SPM, IXPM, and IFWM can be compensated well using digital nonlinearity compensation (NLC) techniques. In WDM superchannel systems, a significant portion of the nonlinear distortion comes from the nonlinear interaction between the channel under consideration and the co-propagating signals in adjacent channels, referred to as inter-channel nonlinearity effects [2]. The inter-channel nonlinear effects can be classified into three types: (a) XPM, (b) cross-polarization modulation (XPolM), and (c) FWM [22, 23]. In recent years, the spectral efficiency of the polarization division multiplexed (PDM) optical communication systems have dramatically improved by introducing a promising detection technique referred to as coherent detection. The coherent detection also enables the implementation of the advanced forward error-correction coding techniques and the adaptive digital signal processing (DSP) algorithms to combat time-varying transmission impairments [24]. A plethora of NLC techniques based on DSP has been developed in the literature to deal with the Kerr nonlinearity effect.

In this tutorial, we provide a comprehensive description of the characteristics of the optical fiber channel, the theory behind the origin of the Kerr effect, the development of the pulse propagation equation, the numerical and analytical tools used for solving the pulse propagation equation, and various Kerr-induced nonlinearity effects in the optical fiber channel. We also provide a concise review of various state-of-the-art DSP techniques available in the literature to deal with the detrimental effects of the Kerr-induced fiber nonlinearity effects. For the maximum benefit of the readers, we adopt a didactic approach with the help of a very detailed mathematical interpretation for various optical phenomenons, processes, models, and methods covered in this tutorial.

The remainder of this tutorial is organized as follows. Section 2 explains the characteristics of the optical fiber channel. Section 3 describes the nonlinear susceptibility and the Kerr effect. Section 4 provides the physics behind the pulse propagation in optical fiber. Section 5 demonstrates various mathematical tools used to solve the pulse propagation equation. Section 6 details various Kerr-induced fiber nonlinearity effects in the optical fiber. Section 7 outlines various nonlinearity compensation techniques used to combat the Kerr-induced nonlinearity effects in the optical fiber. Section 8 illustrates the numerical simulation results and the computational complexity evaluation of the selected nonlinearity compensation techniques. Finally, section 9 concludes the tutorial.

2. Characteristics of optical fiber channel

The optical fiber consists of a glass core surrounded by a cladding layer to confine the modulated light inside the core region. The refractive index n_2 of the cladding layer is slightly lower than that of the core index n_1 to facilitate the total internal reflection [1]. Such fibers are generally referred to as step-index fiber. The step-index optical fibers can be categorized by two parameters, namely core-cladding index difference Δ and the V parameter, given as:

$$\Delta = \frac{n_1 - n_2}{n_1} \tag{1}$$

and

$$V = k_0 a \sqrt{n_1^2 - n_2^2}, \tag{2}$$

$$V = k_0 a \mathrm{NA}, \tag{3}$$

respectively, where $k_0 = \frac{2\pi}{\lambda}$, λ is the wavelength of light, a is the core radius, and $\mathrm{NA} = \sqrt{n_1^2 - n_2^2}$ is the numerical aperture. The V parameter determines the number of propagation modes in the optical fiber. For example, if $V < 2.405$, then the step-index fiber supports only one mode. Such fibers are termed as single-mode fibers [1]. It is important to mention that, in this tutorial, we consider nonlinearity effects in the single-mode fibers since they are used to realize the long-haul high data rate optical communication systems.

Example 1 . The V parameter of a step-index fiber is given by 2.3 at 1550 nm wavelength. Calculate the NA for a core radius of 5 μm.

Solution. From (3),

$$\mathrm{NA} = V\frac{\lambda}{2\pi a} = 2.3 \times \left(\frac{1550\ \mathrm{nm}}{2\pi \times 5\ \mu\mathrm{m}}\right) = 0.11.$$

2.1. Fiber losses

While the optical signal propagates through the optical fiber channel, the fiber losses attenuate the signal power as a function of the transmission distance. The transmitted signal power considering the fiber attenuation can be represented as [1]:

$$P_\mathrm{T} = P_0 \exp(-\alpha L), \tag{4}$$

where P_0 is the power launched at the input of the fiber, α is the fiber attenuation parameter, and L is the transmission distance. It is worth mentioning that the fiber attenuation parameter is usually expressed in units of dB km^{-1}, which can be represented by using (4), as follows:

$$\alpha_{\mathrm{dB\,km^{-1}}} = -\frac{10}{L}\log_{10}\left(\frac{P_\mathrm{T}}{P_0}\right) = 10\log_{10}(e)\,\alpha = 4.343\alpha. \tag{5}$$

The attenuation in optical fiber is caused by several mechanisms, including absorption, scattering, and geometric effects [1]. The material impurity in the silica core fiber causes the absorption of the light energy. The hydroxyl ion (OH) absorption is one of the main absorptions in the case of glass fibers [1]. The OH absorption causes the multiple absorption peaks in the wavelength range from visible to the infrared band [25]. The main source of the scattering loss in optical fiber is due to the Rayleigh scattering [25]. During the fabrication of the optical fiber, the variation of the refractive index is caused by the microscopic variations of fiber material component density, randomly distributed material defects, and inhomogeneous material structure [25]. The scale of this index variation is much smaller than the wavelength of interest. The energy scattering when the propagating light interacts with such small index variation causes the Rayleigh scattering [25]. The significance of the Rayleigh scattering in optical fiber is reduced as the wavelength increases.

The physical bending of the fiber is the main source of the geometric effect causing the signal power attenuation. There are two types of bending loss, including macroscopic and microscopic loss. The macroscopic bending loss is produced whenever the optical fiber is subjected to a significant amount of bending beyond a critical value of curvature [1, 2]. The microscopic bending loss is comparatively weaker and is caused by the strain or stress distributed along the length of the fiber [1, 2].

Example 2 . A single-mode optical fiber exhibits an attenuation of 0.5 dB km^{-1} and 0.22 dB km^{-1} at wavelengths 1310 nm and 1550 nm, respectively. Assuming that two optical signals, one at 1310 nm with an optical power of 120 μW and the other at 1550 nm with an optical power of 80 μW are transmitted through the given fiber. Calculate the power levels in μW of these two signals at transmission distances: (a) 15 km and (b) 80 km.

Solution. Since the attenuation values are given in dB km^{-1}, we need to convert the given power levels from μW to dBm, as follows:

$$P[80\ \mu\text{W}] = 10\log\left(\frac{80\ \mu\text{W}}{1\ \text{mW}}\right) = 10\log(0.08)$$
$$= -10.97\ \text{dBm}$$

$$P[120\ \mu\text{W}] = 10\log\left(\frac{120\ \mu\text{W}}{1\ \text{mW}}\right) = 10\log(0.12)$$
$$= -9.2\ \text{dBm}.$$

(a) At 15 km, we have the following power levels:

$$P_{1310}[15\ \text{km}] = -9.2\ \text{dBm} - \left(0.5\ \text{dB km}^{-1} \times 15\ \text{km}\right)$$
$$= -16.7\ \text{dB km}^{-1} = 21.34\ \mu\text{W}$$

$$P_{1550}[15\ \text{km}] = -10.97\ \text{dBm} - \left(0.22\ \text{dB km}^{-1} \times 15\ \text{km}\right)$$
$$= -14.27\ \text{dB km}^{-1} = 37.41\ \mu\text{W}.$$

(b) At 80 km, we have the following power levels:

$$P_{1310}[80\ \text{km}] = -9.2\ \text{dBm} - \left(0.5\ \text{dB km}^{-1} \times 80\ \text{km}\right)$$
$$= -49.2\ \text{dB km}^{-1} = 0.01\ \mu\text{W}$$

$$P_{1550}[80\ \text{km}] = -10.97\ \text{dBm} - \left(0.22\ \text{dB km}^{-1} \times 80\ \text{km}\right)$$
$$= -28.57\ \text{dB km}^{-1} = 1.4\ \mu\text{W}.$$

2.2. Chromatic dispersion

In general, the response of a medium to the incident electromagnetic wave depends on the optical frequency of ω. This property of the medium is referred to as the CD [1]. The CD has two parts: material dispersion and the waveguide dispersion [26]. The material dispersion part of CD effect in optical fiber manifests through the frequency dependence of the refractive index $n(w)$. In the case of the short optical pulses used in the optical communication system, the material dispersion plays a critical role since the optical pulse consists of several frequency components. In the presence of material dispersion, the different spectral components associated with the pulse travel at different speeds given by $c/n(w)$, where c is the speed of light. That will induce pulse broadening [1]. In the weakly nonlinear regime, the dispersion-induced pulse broadening is the dominant impairment and causes severe penalties to the optical communication systems. In the presence of strong nonlinearity, the interplay between dispersion and nonlinearity is quite complicated, which results in quantitatively different behaviour for the pulse distortions in the optical fiber channel [1].

The effect of dispersion can be quantitatively accounted by expanding the mode-propagation constant β in a Taylor series about the frequency w_0 at which the pulse spectrum is centered as [1]:

$$\beta(w) = n(w)\frac{w}{c} = \beta_0 + \beta_1(w - w_0) + \frac{1}{2}\beta_2(w - w_0)^2 + \cdots, \tag{6}$$

where

$$\beta_m = \left(\frac{d^m\beta}{dw^m}\right)\bigg|_{w=w_0} \qquad m = 1, 2, \ldots, \infty. \tag{7}$$

From (6), β_2 can be represented as:

$$\beta_2 = \frac{1}{c}\left(2\frac{dn}{dw} + w\frac{d^2n}{dw^2}\right). \tag{8}$$

In reality, the envelope of the optical pulse moves at the group velocity, and the parameter β_2 causes the dispersion of the group velocity, which leads to the pulse broadening [1]. This effect is called the group-velocity dispersion (GVD), and β_2 is termed as the GVD parameter [1].

In the optical fiber, a part of the optical signal propagates through the cladding, referred to as the dielectric waveguiding, which causes a slight reduction in the effective mode index $n(w)$ of the core. That results in the waveguide dispersion and must be added to the material dispersion [26]. In general, the contribution of the waveguide dispersion is relatively small when compared to the material dispersion except near the zero-dispersion wavelength λ_0. For standard fibers, the effect of the waveguide dispersion is to shift λ_0 slightly towards the longer wavelengths, such as $\lambda_0 \approx 1.31\ \mu$m. It is important to mention that the quantity dispersion parameter D is related to β_2 as [26]:

$$D = -\frac{2\pi c}{\lambda^2}\beta_2 \approx \frac{\lambda d^2 n}{c d\lambda^2}. \tag{9}$$

Example 3 . An optical signal of wavelength 1550 nm is transmitted trough a single-mode fiber having a GVD of $-25.509 \times 10^3\ \text{fs}^2\, m^{-1}$. Calculate the dispersion parameter.

Solution. Using (9),

$$D = -\frac{2\pi c}{\lambda^2}\beta_2$$
$$= -\left(\frac{2\pi \times 3 \times 10^8\ \text{m s}^{-1}}{1550^2\ \text{nm}^2}\right) \times \left(-25.509 \times 10^3\ \text{fs}^2\,\text{m}^{-1}\right)$$
$$= 20\ \text{ps}\,(\text{nm} \cdot \text{km})^{-1}.$$

It is worth mentioning that the waveguide dispersion is dependent on the fiber design parameters such as the core radius a, and the core-cladding index difference Δ. The zero-dispersion wavelength λ_0 can be shifted to the vicinity of 1.55 μm by using this feature of the waveguide dispersion. It is also worthy of mentioning that the fiber loss is minimum at 1.55 μm wavelength range [26]. Such dispersion-shifted fibers are commonly used in communication systems.

The nonlinearity in optical fiber exhibits a different response behaviour depending on the sign of the GVD parameter. For example, for wavelengths $\lambda < \lambda_0$, the fiber exhibits normal dispersion regime where $\beta_2 > 0$. In this regime, the low-frequency components travel faster than the high-frequency components of the same optical pulse. The fiber is said to be in an anomalous dispersion regime when $\beta_2 < 0$. In silica-based optical fibers, the anomalous dispersion regime occurs when the light wavelength exceeds the zero-dispersion

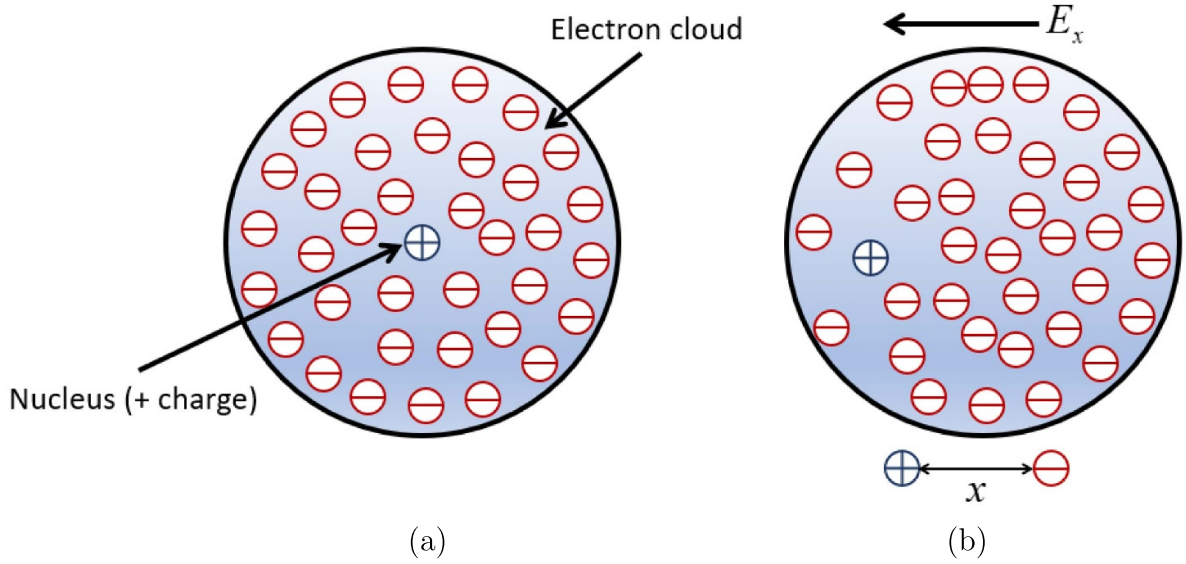

Figure 1. Classical electron oscillator model. (a) in equilibrium and (b) in the presence of an external field. © 2014 John Wiley & Sons, Ltd Reprinted, with permission, from [27].

wavelength, i.e. $\lambda > \lambda_0$. The anomalous dispersion regime is of particular interest in optical communication systems since it supports the existence of solitons through a balance between the dispersive and nonlinear effects [1].

The CD-induced mismatch in the group velocities of the optical pulses at different wavelengths causes a different speed for those pulses in the optical fiber [1]. This feature leads to an important effect referred to as the walk-off effect [1]. More specifically, when the fast-moving pulse completely walks through the slower moving pulse, the nonlinear interaction between two optical pulses ceases to occur [1]. That is governed by a walk-off parameter defined as [1]:

$$d_{12} = \beta_1(\lambda_1) - \beta_1(\lambda_2) = v_g^{-1}(\lambda_1) - v_g^{-1}(\lambda_2), \tag{10}$$

where λ_1 and λ_2 are the centre wavelengths of the two pulses and $v_g = \frac{1}{\beta_1}$ is the group velocity and is calculated using (7). For pulses of width τ, the walk-off length can be defined as [1]:

$$L_w = \frac{\tau}{|d_{12}|}. \tag{11}$$

3. Nonlinear susceptibility and the Kerr effect in optical fibers

3.1. Nonlinear susceptibility

For intense electromagnetic fields, the response of the optical fiber is nonlinear. The origin of the nonlinear response is due to the anharmonic motion of bound electrons under the influence of the applied electromagnetic field [27]. The electric field of the incident light interacts with the electron and makes it oscillate in accordance with Coulomb's law [27]. The oscillating charge resembles an antenna and radiates the electromagnetic energy at the same frequency as the incident field with a different phase shift. The dynamics of the displaced electron under the influence of the applied electric field is a fundamental field of study in quantum mechanics [27].

On the other hand, in the classical electron oscillator model, the electron is modelled as a charged cloud surrounding the nucleus, as shown in figure 1(a). Figure 1(b) shows that the electron charge cloud is displaced when an electric field E_x is applied [27]. According to Newton's law, the equation of motion for the center of the electron charge cloud can be represented as [27]:

$$m\frac{d^2x}{dt^2} = F_{\text{ext}} = q_e E_x, \tag{12}$$

where m is the electron mass, $x(t)$ is the displacement, and q_e is the electron charge. There exists a force of attraction between the nucleus and the electron charge cloud when the electron charge cloud moves away for the equilibrium position [27]. For a small displacement $x(t)$, the restoration force can be approximated as [27]:

$$F_{\text{restoration}} = -Kx, \tag{13}$$

where K is a constant.

The negative sign in (13) indicates that the restoration force acts in a direction opposite to the external force [27]. This situation is similar to the case of a simple pendulum pushed away from the equilibrium position by an external force [27]. The restoration force due to the gravitation pulls back the pendulum to the equilibrium position. The net force acting on the electron can be represented as [27]:

$$F_{\text{net}} = F_{\text{ext}} + F_{\text{restoration}} = q_e E_x - Kx. \tag{14}$$

Next, by combining (12) and (14), we obtain [27]:

$$m\frac{d^2x}{dt^2} = F_{\text{net}} = q_e E_x - Kx \tag{15}$$

or

$$\frac{d^2x}{dt^2} + w_0^2 x = \left(\frac{q_e}{m}\right) E_x, \tag{16}$$

where $w_0 = \sqrt{K/m}$ is the natural frequency of oscillation. Assume that the applied electric field is of the form as given below [27]:

$$E_x = E_0 \exp(-jwt). \tag{17}$$

The displacement $x(t)$ due to the applied field also change harmonically in the steady state and try a trial solution as [27]:

$$x(t) = B \exp(-jwt). \tag{18}$$

Substituting (17) and (18) in (16), we get [27]:

$$B = \frac{E_0 q_e}{m(w_0^2 - w^2)}, \tag{19}$$

and

$$x(t) = \frac{q_e}{m(w_0^2 - w^2)} E_x. \tag{20}$$

The dipole moment of an atom is given as [27]:

$$p_x = q_e x(t) = \frac{q_e^2}{m(w_0^2 - w^2)} E_x. \tag{21}$$

In general [27]:

$$\mathbf{p} = \frac{q_e^2}{m(w_0^2 - w^2)} \mathbf{E}. \tag{22}$$

If we define $\mathbf{P} = N\mathbf{p}$ as the dielectric polarization, (22) can be written as [27]:

$$\mathbf{P} = \frac{Nq_e^2}{m(w_0^2 - w^2)} \mathbf{E} \tag{23}$$

where N is the number of atoms per unit volume.

For the weak incident electromagnetic field, we can relate the dielectric polarization $\mathbf{P}$ to the electric field intensity $\mathbf{E}$ as [27]:

$$\mathbf{P} = \varepsilon_0 \chi^{(1)} \mathbf{E}, \tag{24}$$

where ε_0 is the electric permittivity of free space and $\varepsilon_0\chi^{(1)}$ is the first-order susceptibility or linear susceptibility [27], and:

$$\chi^{(1)} = \frac{Nq_e^2}{m(w_0^2 - w^2)\varepsilon_0}. \tag{25}$$

It is important to mention that if the medium is not isotropic, the susceptibility depends on direction as well and (24) is modified as [27]:

$$P_j = \chi_{jx}^{(1)} E_x + \chi_{jy}^{(1)} E_y + \chi_{jz}^{(1)} E_z, \quad j = x, y, z \tag{26}$$

or

$$\mathbf{P} = \varepsilon_0 \chi^{(1)}.\mathbf{E}, \tag{27}$$

where $\chi^{(1)}$ is a 3×3 matrix and $\cdot$ denotes dot product.

If the incident electromagnetic field is intense, the relation between the restoration force and the displacement is nonlinear, and thereby, the electron cloud oscillation is not harmonic [27]. In this case, the relation between the dielectric polarization $\mathbf{P}$ and the electric dipoles are nonlinear, which can be generalized as [27]:

$$\mathbf{P} = \varepsilon_0 \left(\chi^{(1)} \cdot \mathbf{E} + \chi^{(2)} : \mathbf{EE} + \chi^{(3)} \vdots \mathbf{EEE} + \ldots \right), \tag{28}$$

where $\chi^{(j)}$ $(j = 1, 2, \ldots)$ is the jth order susceptibility. $\chi^{(j)}$ is also a tensor of rank $j+1$. $\chi^{(1)}$ is the first-order susceptibility and is related to the linear refractive index. The second-order susceptibility $\chi^{(2)}$ generates the second-harmonic and the sum-frequency terms [27]. However, since SiO_2 is a symmetric molecule, $\chi^{(2)}$ vanishes for the silica glasses [27]. As a result, the optical fibers do not exhibit the $\chi^{(2)}$-induced nonlinear effects [27].

3.2. The Kerr effect

The nonlinearity effects in optical fiber originate from the third-order susceptibility $\chi^{(3)}$. One of the primary sources of the nonlinearity effect in optical fiber is the $\chi^{(3)}$-induced nonlinear refraction, the Kerr effect, a phenomenon referring to the light intensity-dependent refractive index [27]. Assume that the electromagnetic field incident on the optical fiber core has only E_x and H_y components. Then, the tensor equation in (28) can be simplified for a centrally symmetric dielectric material as [27]:

$$P_x = \varepsilon_0 \chi_{xx}^{(1)} E_x + \varepsilon_0 \chi_{xxxx}^{(3)} E_x^3, \tag{29}$$

where $\chi_{xxxx}^{(3)}$ is a component of the fourth-rank tensor $\chi^{(3)}$. Suppose, the incident optical field is a monochromatic wave given as:

$$E_x = E_0 \exp(-jwt). \tag{30}$$

To find E_x^3, we should first find the real part of E_x, i.e.

$$\mathrm{Re}[E_x] = \frac{1}{2}\left[E_0 \exp(-jwt) + E_0^* \exp(jwt)\right], \tag{31}$$

$$\{\mathrm{Re}[E_x]\}^3 = \frac{1}{8}\left\{E_0^3 \exp(-j3wt) + E_0^{*3} \exp(j3wt) + 3|E_0|^2 \left[E_0 \exp(-jwt) + E_0^* \exp(jwt)\right]\right\}. \tag{32}$$

In the absence of the special phase-matching techniques, the third harmonic terms in (32) can be neglected [27].

Let the polarization at frequency w be:

$$P_x = P_0 \exp(-jwt), \tag{33}$$

Then,

$$\mathrm{Re}[P_x] = \frac{1}{2}\left[P_0 \exp(-jwt) + P_0^* \exp(jwt)\right]. \tag{34}$$

From (29), we have:

$$\mathrm{Re}[P_x] = \varepsilon_0 \chi_{xx}^{(1)} \mathrm{Re}[E_x] + \varepsilon_0 \chi_{xxxx}^{(3)} \mathrm{Re}[E_x]^3, \tag{35}$$

where the imaginary parts of the susceptibility are ignored. Substituting (31) and (32) into (35), collecting the terms that are proportional to $\exp(-jwt)$, and comparing it with (34), we obtain:

$$P_0 = \varepsilon_0 \left(\chi_{xx}^{(1)} + \frac{3|E_0|^2}{4}\chi_{xxxx}^{(3)}\right) E_0 = \varepsilon_0 \chi_{\mathrm{eff}} E_0, \tag{36}$$

where χ_{eff} is the effective susceptibility that includes both linear and nonlinear susceptibilities [27].

We can represent the electric field density D as [27]:

$$D = \varepsilon_0 E_0 + P_0. \tag{37}$$

Substituting (36) in (37), we can write:

$$D = \varepsilon_0 \left[1 + \chi_{xx}^{(1)} + \frac{3|E_0|^2}{4}\chi_{xxxx}^{(3)}\right] E_0. \tag{38}$$

In general, we can represent the electric flux density as [27]:

$$D = \varepsilon_0 \varepsilon_r E_0, \tag{39}$$

where ε_r is the relative permittivity. From (38) and (39), we can represent ε_{r} as:

$$\varepsilon_{\mathrm{r}} = 1 + \chi_{xx}^{(1)} + \frac{3|E_0|^2}{4}\chi_{xxxx}^{(3)}. \tag{40}$$

Since the relative permittivity ε_{r} and refractive index n are related by $n^2 = \varepsilon_{\mathrm{r}}$, we can write:

$$\begin{aligned} n^2 &= 1 + \chi_{xx}^{(1)} + \frac{3|E_0|^2}{4}\chi_{xxxx}^{(3)} \\ &= n_0^2 + \frac{3|E_0|^2}{4}\chi_{xxxx}^{(3)}, \end{aligned} \tag{41}$$

where n_0 is the linear refractive index and the second term of (41) represents the nonlinear contribution to the refractive index.

From (41), we can represent:

$$\begin{aligned} n &= n_0 \left(1 + \frac{3|E_0|^2}{4n_0^2}\chi_{xxxx}^{(3)}\right)^{1/2} \\ &\cong n_0 + n_2 |E_0|^2, \end{aligned} \tag{42}$$

where

$$n_2 = \frac{3\chi_{xxxx}^{(3)}}{8n_0}. \tag{43}$$

In (43), the term n_2 is called the Kerr coefficient [27]. For silica based fiber, the typical value of n_2 varies between 1.2×10^{-20} m^2 W^{-1} and 3.2×10^{-20} m^2 W^{-1} [27]. From (42), it is clear that the nonlinear part of the refractive index n is proportional to the optical intensity $|E_0|^2$. This effect is termed as the Kerr effect [27].

4. Pulse propagation in optical fibers

The wave propagation in dispersive nonlinear media is governed by the fundamental theory of electromagnetic wave propagation underpinned by the Maxwell's equations [1]. Using Maxwell's equations, one can easily show that:

$$\nabla \times \nabla \times \mathbf{E} = -\frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} - \mu_0 \frac{\partial^2 \mathbf{P}}{\partial t^2}, \tag{44}$$

where μ_0 is the permeability of the free space.

For the Kerr effect-based nonlinearity in optical fiber, the induced dielectric polarization $\mathbf{P}(\mathbf{r},t)$ consists of two parts such as [1]:

$$\mathbf{P}(\mathbf{r},t) = \mathbf{P}_L(\mathbf{r},t) + \mathbf{P}_{\mathrm{NL}}(\mathbf{r},t), \tag{45}$$

where $\mathbf{P}_{\mathrm{L}}$ and $\mathbf{P}_{\mathrm{NL}}$ are the linear and nonlinear parts which are given as:

$$\mathbf{P}_{\mathrm{L}}(\mathbf{r},t) = \varepsilon_0 \int_{-\infty}^{\infty} \chi^{(1)}(t-t').\mathbf{E}(\mathbf{r},t')\, dt', \tag{46}$$

and

$$\begin{aligned} \mathbf{P}_{\mathrm{NL}}(\mathbf{r},t) = \varepsilon_0 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \chi^{(3)}(t-t_1,t-t_2,t-t_3) \vdots \mathbf{E}(\mathbf{r},t_1) \\ \times \mathbf{E}(\mathbf{r},t_2)\mathbf{E}(\mathbf{r},t_3)\, dt_1 dt_2 dt_3. \end{aligned} \tag{47}$$

4.1. Nonlinear Schrödinger equation

From (44) and (45), we can represent the wave equation as:

$$\nabla^2 \mathbf{E} - \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} = \mu_0 \frac{\partial^2 \mathbf{P}_{\mathrm{L}}}{\partial t^2} + \mu_0 \frac{\partial^2 \mathbf{P}_{\mathrm{NL}}}{\partial t^2}, \tag{48}$$

where $\mathbf{P}_{\mathrm{L}}$ and $\mathbf{P}_{\mathrm{NL}}$ are given by (46) and (47), respectively. To solve the wave equation in (48), we adopt several simplifying assumptions including [1]:

(a) The nonlinear part of the dielectric polarization $\mathbf{P}_{\mathrm{NL}}$ is treated as a small perturbation to the linear part $\mathbf{P}_{\mathrm{L}}$;
(b) A scalar approach is adopted, i.e. the optical field is assumed to maintain the polarization along the length of the fiber; and
(c) A quasi-monochromatic assumption for the optical field.

By adopting the slowly varying envelop approximation, we can represent the electric field $\mathbf{E}(\mathbf{r},t)$ as [1]:

$$\mathbf{E}(\mathbf{r},t) = \frac{1}{2}\hat{x}\left[E_0(\mathbf{r},t)\exp(-jw_0 t) + E_0^*(\mathbf{r},t)\exp(jw_0 t)\right], \tag{49}$$

where $\hat{x}$ is the polarization unit vector and w_0 is the center frequency of the optical pulse spectrum. The dielectric polarization components $\mathbf{P}_{\mathrm{L}}$ and $\mathbf{P}_{\mathrm{NL}}$ can also be expressed in similar form as [1]:

$$\mathbf{P}_{\mathrm{L}}(\mathbf{r},t) = \frac{1}{2}\hat{x}\left[P_{\mathrm{L}}(\mathbf{r},t)\exp\left(-jw_0t\right) + P^{*}_{\mathrm{L}}(\mathbf{r},t)\exp\left(jw_0t\right)\right], \tag{50}$$

and

$$\mathbf{P}_{\mathrm{NL}}(\mathbf{r},t) = \frac{1}{2}\hat{x}\left[P_{\mathrm{NL}}(\mathbf{r},t)\exp\left(-jw_0t\right) + P^{*}_{\mathrm{NL}}(\mathbf{r},t)\exp\left(jw_0t\right)\right]. \tag{51}$$

The linear component $\mathbf{P}_{\mathrm{L}}$ can be calculated by substituting (50) in (46) as [1]:

$$\begin{aligned}\mathbf{P}_{\mathrm{L}}(\mathbf{r},t) &= \varepsilon_0 \int_{-\infty}^{\infty} \chi^{(1)}_{xx}(t-t')E_0(\mathbf{r},t')\exp\left(jw_0(t-t')\right)\,dt' \\ &= \frac{\varepsilon_0}{2\pi}\int_{-\infty}^{\infty} \tilde{\chi}^{(1)}_{xx}(w)\tilde{E}_0(\mathbf{r},w-w_0)\exp\left(-j(w-w_0)t\right)\,dw,\end{aligned} \tag{52}$$

where $\tilde{E}_0(\mathbf{r},w)$ is the Fourier transform of $E_0(\mathbf{r},w)$.

The nonlinear component $\mathbf{P}_{\mathrm{NL}}$ is obtained by substituting (51) in (47) and after some simplifications we obtain:

$$\mathbf{P}_{\mathrm{NL}}(\mathbf{r},t) = \varepsilon_0\chi^{(3)}\vdots\mathbf{E}(\mathbf{r},t)\mathbf{E}(\mathbf{r},t)\mathbf{E}(\mathbf{r},t). \tag{53}$$

When (49) is substituted in (53) and following the analysis given in section 3.1, we can approximate $\mathbf{P}_{\mathrm{NL}}$ as [1]:

$$\mathbf{P}_{\mathrm{NL}}(\mathbf{r},t) \approx \varepsilon_0\varepsilon_{\mathrm{NL}}E_0(\mathbf{r},t), \tag{54}$$

where

$$\varepsilon_{\mathbf{NL}} = \frac{3\chi^{(3)}_{xxxx}}{4}\left|E_0(\mathbf{r},t)\right|^2. \tag{55}$$

For simplicity, we adopt a frequency-domain analysis to derive the wave equation for the slowly varying amplitude $E_0(\mathbf{r},t)$. Substituting (49)–(51) in (48), the Fourier transform $\tilde{E}_0(\mathbf{r},w-w_0)$ can be defined as:

$$\tilde{E}_0(\mathbf{r},w-w_0) = \int_{-\infty}^{\infty} E_0(\mathbf{r},t)\exp\left(j(w-w_0)t\right)\,dt, \tag{56}$$

which is found to satisfy the Helmholtz equation as given below [1]:

$$\nabla^2\tilde{E} + \varepsilon(w)k_0^2\tilde{E} = 0, \tag{57}$$

where $k_0 = \frac{w}{c}$ and

$$\varepsilon(w) = 1 + \tilde{\chi}^{(1)}_{xx}(w) + \varepsilon_{\mathrm{NL}}. \tag{58}$$

Equation (57) can be solved using the method of separation of variables. Assuming the solution of the form:

$$\tilde{E}_0(\mathbf{r},w-w_0) = F(x,y)Q(z,w-w_0)\exp\left(j\beta_0z\right), \tag{59}$$

where $Q(z,w)$ is a slowly varying function of z and β_0 is the wave number. From (57), we can write two equations for $F(x,y)$ and $Q(z,w)$ as:

$$\frac{\partial^2F}{\partial x^2} + \frac{\partial^2F}{\partial y^2} + \left[\varepsilon(w)k_0^2 - \tilde{\beta}^2\right]F = 0, \tag{60}$$

and

$$2j\beta_0\frac{\partial Q}{\partial z} + \left(\tilde{\beta}^2 - \beta_0^2\right)Q = 0. \tag{61}$$

The dielectric constant $\varepsilon(w)$ can be approximated as:

$$\varepsilon = (n_0 + \Delta n)^2 \approx n_0^2 + 2n_0\Delta n, \tag{62}$$

where Δn is a small perturbation given by:

$$\Delta n = n_2\left|E_0\right|^2 + \frac{j\tilde{\alpha}}{2k_0}, \tag{63}$$

where $\tilde{\alpha}$ is the absorption coefficient.

The first-order perturbation theory can be used to solve (60). After solving using the first-order perturbation theory, the value of $\tilde{\beta}$ can be represented as [1]:

$$\tilde{\beta}(w) = \beta(w) + \Delta\beta, \tag{64}$$

where

$$\Delta\beta = \frac{k_0\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Delta n\left|F(x,y)\right|^2\,dxdy}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left|F(x,y)\right|^2\,dxdy}. \tag{65}$$

Similarly, the electric field $\mathbf{E}(\mathbf{r},t)$ can be written as [1]:

$$\begin{aligned}\mathbf{E}(\mathbf{r},t) = \frac{1}{2}\hat{x}\left[F(x,y)q(z,t)\exp\left(j(\beta_0z - w_0t)\right)\right. \\ \left.+F^{*}(x,y)q^{*}(z,t)\exp\left(-j(\beta_0z - w_0t)\right)\right],\end{aligned} \tag{66}$$

where $q(z,t)$ is the slowly varying pulse envelop [1].

The Fourier transform $Q(z,w-w_0)$ of $q(z,t)$ satisfies (61), which can be represented as [1]:

$$\frac{\partial Q}{\partial z} = j\left[\beta(w) + \Delta\beta - \beta_0\right]Q. \tag{67}$$

The propagation equation for $q(z,t)$ is obtained by taking the inverse Fourier transform of (67). In (67), the exact functional form of the mode-propagation constant $\beta(w)$ is rarely known and therefore, it is useful to expand $\beta(w)$ in Taylor series as in (6) [1]. Because of the quasi-monochromatic assumption for the optical field, the cubic and higher-order terms in (6) can be neglected [1]. Next, substitute (6) in (67) and take the inverse Fourier transform, we obtain:

$$\frac{\partial}{\partial z}q(z,t) = -j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}q(z,t) + j\Delta\beta q(z,t). \tag{68}$$

The term $\Delta\beta$ in (68) includes the effect of loss and the nonlinearity [1]. Next, evaluate $\Delta\beta$ using (63) and (65), and substitute in (68), we obtain the nonlinear Schrödinger equation (NLSE) as:

$$\frac{\partial}{\partial z}q(z,t)+\frac{\alpha}{2}q(z,t)+j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}q(z,t)=j\gamma\left|q(z,t)\right|^2 q(z,t), \tag{69}$$

where z is the propagation distance, β_2 is the GVD parameter, α is the signal attenuation parameter, and γ is the nonlinearity coefficient given as [1]:

$$\gamma=\frac{n_2 w_0}{cA_{\text{eff}}}, \tag{70}$$

where A_{eff} is the effective area of the fiber and is given as:

$$A_{\text{eff}}=\frac{\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left|F(x,y)\right|^2 dxdy\right)^2}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left|F(x,y)\right|^4 dxdy}. \tag{71}$$

Using (70), one can calculate the peak nonlinear phase shift accumulated over the link as:

$$\varphi_{\text{NL}}=\gamma L_{\text{eff}}P_0, \tag{72}$$

where P_0 is the peak power and $L_{\text{eff}}=\frac{1-\exp(-\alpha L)}{\alpha}$, with L as the fiber span length.

Example 4 . For a given single-mode fiber with Kerr coefficient $n_2=2.6\times10^{-20}\,\text{m}^2\,\text{W}^{-1}$ and the effective area $A_{\text{eff}}=80\,\mu\text{m}^2$, calculate the nonlinearity coefficient γ at a wavelength of 1550 nm.

Solution.

$$\begin{aligned}\gamma&=\frac{n_2 w_0}{cA_{\text{eff}}}=\frac{n_2 2\pi f_0}{cA_{\text{eff}}}=\frac{2\pi n_2}{\lambda_0 A_{\text{eff}}}\\&=\frac{2\pi\times2.6\times10^{-20}\,\text{m}^2\,\text{W}^{-1}}{1550\,\text{nm}\times80\,\mu\text{m}^2}\\&=1.32\times10^{-3}\,\text{W}^{-1}\text{m}^{-1}.\end{aligned}$$

The evaluation of A_{eff} requires the use of modal distribution $F(x,y)$ for the fundamental fiber mode [1]. It is worth mentioning that the typical value of A_{eff} varies in the range $20-100\,\mu\text{m}^2$ in the 1550 nm region depending on the fiber design [1]. As a result, the nonlinearity coefficient γ takes values in the range $1-10\,\text{W}^{-1}\,\text{km}^{-1}$ if $n_2\approx2.6\times10^{-20}\,\text{m}^2\,\text{W}^{-1}$ [1].

The NLSE in (69) can be simplified by applying the transformation $q(z,t)\triangleq u(z,t)\exp(-\frac{\alpha}{2}z)$, where $u(z,t)$ is the normalized field, and the simplified form can be written as [1]:

$$\frac{\partial}{\partial z}u(z,t)+j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}u(z,t)=j\gamma\left|u(z,t)\right|^2 u(z,t)\exp(-\alpha z). \tag{73}$$

Example 5 . The long-haul optical fiber transmission system has following parameters: attenuation coefficient $\alpha=0.2\,\text{dB}\,\text{km}^{-1}$, fiber span length $L=80$ km, Kerr coefficient $n_2=2.6\times10^{-20}\,\text{m}^2\,\text{W}^{-1}$, wavelength $\lambda_0=1550$ nm, and the peak power $P_0=2$ dBm. For a transmission distance of 1200 km, calculate lower limit on the effective area A_{eff} of the fiber in which the peak nonlinear phase shift φ_{NL} accumulated over the link should be less than 0.5 rad. Ignore β_2.

Solution. The first step is to convert α from $\text{dB}\,\text{km}^{-1}$ to 1/km as follows:

$$\begin{aligned}\alpha_{1/\text{km}}&=\alpha_{\text{dB}\,\text{km}^{-1}}\left(\frac{\ln(10)}{10}\right)\\&=0.2\,\text{dB}\,\text{km}^{-1}\times0.23\\&=0.046\,1/\text{km}\\L_{\text{eff}}&=\frac{1-\exp(-\alpha L)}{\alpha}\\&=\frac{1-\exp(-0.046\,1/\text{km}\times80\,\text{km})}{0.046\,1/\text{km}}\\&=21.2\,\text{km}.\end{aligned}$$

The number of fiber spans is given by:

$$\frac{1200\,\text{km}}{80\,\text{km}}=15.$$

The total nonlinear phase shift $=15\varphi_{\text{NL}}$.

$$P_0=10^{2\ \text{dBm}/10}\,\text{mW}=1.58\,\text{mW}.$$

As per the requirement $15\varphi_{\text{NL}}<0.5$.

$$15\varphi_{\text{NL}}<0.5$$

$$15\times\gamma\times21.2\times10^3\times1.58\times10^{-3}<0.5$$

$$\gamma<1\times10^{-3}\,\text{W}^{-1}\text{m}^{-1}$$

$$\gamma=\frac{2\pi n_2}{\lambda_0 A_{\text{eff}}}<1\times10^{-3}\,\text{W}^{-1}\text{m}^{-1}$$

$$A_{\text{eff}}>\frac{2\pi\times2.6\times10^{-20}\,\text{m}^2\,\text{W}^{-1}}{1550\,\text{nm}\times1\times10^{-3}\,\text{W}^{-1}\text{m}^{-1}}$$

$$A_{\text{eff}}>105.34\,\mu\text{m}^2.$$

Therefore, the effective area A_{eff} should be higher than $105.34\,\mu\text{m}^2$ to have the peak nonlinear phase shift φ_{NL} less than or equal to 0.5 rad.

4.2. Manakov equation

For dual-polarization transmission systems, the coupled NLSE (CNLSE) provides an accurate model for the nonlinear pulse propagation in optical fiber [1]. It includes the polarization mode dispersion (PMD) effects in the fiber along with the dispersive and nonlinearity effects. In general, similar to NLSE, the CNLSE also needs to be solved numerically [1]. However, the different length scales associated with PMD, GVD, and nonlinearity effects make the numerical evaluation cumbersome. The dispersive and nonlinearity effects vary on a length

scale from 10 to 100 km, while the birefringence in optical fiber varies on a length scale of 10–100 m [1]. Therefore, the step size used in the numerical evaluation of CNLSE must be less than 1 m [1]. That increases the computation time. Based on this fact, we adopt an approximation method to solve the CNLSE. It is observed that the birefringence fluctuation changes the state of polarization (SOP) of the optical field on a short length scale that the field covers the entire Poincaré sphere after a few kilometers [1]. As a result, the nonlinearity terms in CNLSE can be averaged over the birefringence fluctuations [1]. The resultant propagation equation is referred to as the Manakov equation and can be represented as:

$$\frac{\partial}{\partial z}\mathbf{u}(z,t)+j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}\mathbf{u}(z,t)=j\frac{8}{9}\gamma\left|\mathbf{u}(z,t)\right|^2\mathbf{u}(z,t)\exp\left(-\alpha z\right), \tag{74}$$

where $\mathbf{u}(z,t)=[u_x(z,t)\ \ u_y(z,t)]^{\dagger}$ and the superscript † represents the transpose [1]. It is clear from (74) that the rapid random variations in the SOP of the optical field reduce the effect of the nonlinearity parameter γ by a factor of $\frac{8}{9}$.

For WDM systems, the optical field $\mathbf{u}(z,t)$ in (74) consists of the sum of all the multiplexed channels, which can be represented as:

$$\mathbf{u}(z,t)=\sum_{n'}\mathbf{u}_n(z,t)\exp\left(jw_nt\right). \tag{75}$$

The Manakov equation for the WDM system consists of both intra- and inter-channel nonlinearity effects such as SPM, XPM, and FWM [1]. By substituting (75) in (74), we can represent the propagation equation for the ith channel as:

$$\frac{\partial}{\partial z}u_{x/y,i}=u^{\mathrm{CD}}_{x/y,i}+u^{\mathrm{SPM}}_{x/y,i}+u^{\mathrm{XPM}}_{x/y,m\neq i}+u^{\mathrm{FWM}}_{x/y,k\neq l,m,l\neq i,m\neq i}, \tag{76}$$

where x and y in this context represent the horizontal and vertical polarizations, respectively, $u_{x/y}$ represents the optical field transmitted on x/y polarization tributaries, m, l, and k denote the symbol indices. It is important to note that in (76), the evolution of the optical field envelop of the ith channel along the optical fiber link consists of the CD part arising from the pulse dispersion, the SPM part generating from the nonlinear interaction of the ith channel to itself, XPM term inducing from the nonlinear interaction of the ith channel with one other co-propagating channel, and the FWM term originating from the nonlinear interaction of the ith channel with two or more co-propagating channels.

In (76), the definitions $u^{\mathrm{CD}}_{x/y,i}$, $u^{\mathrm{SPM}}_{x/y,i}$, $u^{\mathrm{XPM}}_{x/y,m\neq i}$, and $u^{\mathrm{FWM}}_{x/y,k\neq l,m,l\neq i,m\neq i}$ can be represented as:

$$u^{\mathrm{CD}}_{x/y,i}=-j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}u_{x/y,i}, \tag{77}$$

$$u^{\mathrm{SPM}}_{x/y,i}=j\frac{8}{9}\gamma\exp\left(-\alpha z\right)\left(\left|u_{x/y,i}\right|^2+\left|u_{y/x,i}\right|^2\right)u_{x/y,i}, \tag{78}$$

$$u^{\mathrm{XPM}}_{x/y,m\neq i}=j\frac{8}{9}\gamma\exp\left(-\alpha z\right)\left(\underbrace{\sum_{m\neq i}\left(2\left|u_{x/y,m}\right|^2+\left|u_{y/x,m}\right|^2\right)u_{x/y,i}}_{\text{Coherent XPM}}+\underbrace{\sum_{m\neq i}\left(u^*_{y/x,m}u_{x/y,m}\right)u_{y/x,i}}_{\text{Incoherent XPM}}\right) \tag{79}$$

$$u^{\mathrm{FWM}}_{x/y,k\neq l,m,l\neq i,m\neq i}=j\frac{8}{9}\gamma\exp\left(-\alpha z\right)\sum_{p',q',r'}\times\sum_{k\neq l,m}\sum_{l\neq i}\sum_{m\neq i}u^*_{p',k}u_{q',l}u_{r',m}, \tag{80}$$

where p',q', and r' takes the x and y polarization tributaries. For simplicity, the space and time variables z,t are not shown in equations (77)–(80).

5. Mathematical tools for solving the propagation equation

5.1. Numerical approximation method

The analytical solution of NLSE in (73) is not available other than for some exceptional cases [1]. Numerical approximations are typically used to solve the propagation equation in (73). The most commonly used numerical method to solve the pulse propagation problem in nonlinear and dispersive media is the split-step Fourier method (SSFM) [1].

5.1.1. The split-step Fourier method. By separating the linear and nonlinear parts, the NLSE in (73) can be written as [1]:

$$\frac{\partial}{\partial z}u(z,t)=\left(\hat{D}+\hat{N}\right)u(z,t) \tag{81}$$

$$\hat{D}=-j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2} \tag{82}$$

$$\hat{N}=j\gamma\left|u(z,t)\right|^2\exp(-\alpha z), \tag{83}$$

where $\hat{D}$ and $\hat{N}$ represent the linear and nonlinear operators, respectively [1].

The SSFM is an iterative signal propagation modelling method in which the optical fiber span is divided into small segments, each having a length of h. More specifically, the signal propagation from z to $z+h$ is carried out in two separate steps [1]. First, the linear operator $\hat{D}$ is set to zero, and only the nonlinearity is taken into account. Second, the nonlinear operator $\hat{N}$ is set to zero, and only the dispersion is taken into account.

Mathematically:

$$u(z+h,t) \approx \exp\left(h\hat{D}\right)\exp\left(h\hat{N}\right)u(z,t). \tag{84}$$

In (84), the step size h is selected in such a way that the linear and the nonlinear sections in each fiber segment can be modelled as independent operations [1].

Example 6 . Show that the SSFM in (84) is accurate to the second-order in h.

Solution. We can start with the Baker–Hausdorff formula [28] for two non-commutating operators μA and μB as:

$$\exp(\mu A)\exp(\mu B) = \exp\left(\mu A + \mu B + \frac{\mu^2}{2}[A,B] + \frac{\mu^3}{12}[A-B,[A,B]] + \cdots\right), \tag{85}$$

where $[A,B] = AB - BA$. By substituting $\mu = h$, $A = \hat{D}$, and $B = \hat{N}$ in (85), we can write:

$$\exp\left(h\hat{D}\right)\exp\left(h\hat{N}\right) = \exp\left(h\hat{D} + h\hat{N} + \frac{h^2}{2}\left[\hat{D},\hat{N}\right] + \frac{h^3}{12}\left[\hat{D}-\hat{N},\left[\hat{D},\hat{N}\right]\right] + \cdots\right). \tag{86}$$

By substituting (86) in (84), we get:

$$u(z+h,t) \approx \exp\left(h\hat{D} + h\hat{N} + \frac{h^2}{2}\left[\hat{D},\hat{N}\right] + \frac{h^3}{12}\left[\hat{D}-\hat{N},\left[\hat{D},\hat{N}\right]\right] + \cdots\right)\exp u(z,t). \tag{87}$$

By examining (87), we understand that the dominant error term is the commutator term $\frac{h^2}{2}\left[\hat{D},\hat{N}\right]$. From this, we can confirm that the SSFM in (84) is accurate to the second-order in h.The accuracy of the SSFM can be improved by writing the symmetric form of exponential operators in (84). This method is referred to as symmetric SSFM. The symmetric SSFM can be represented as follows [1]:

$$u(z+h,t) \approx \exp\left(\frac{h}{2}\hat{D}\right)\exp\left(\int_z^{z+h}\hat{N}(z')dz'\right) \times \exp\left(\frac{h}{2}\hat{D}\right)u(z,t). \tag{88}$$

In this method, the nonlinearity is included in the middle of the fiber segment rather than at the segment boundaries, as shown in figure 2. It is important to mention that for small values of step size h, the term $\exp\left(\int_z^{z+h}\hat{N}(z')dz'\right)$ can be approximated as $\exp\left(h\hat{N}\right)$ [1].

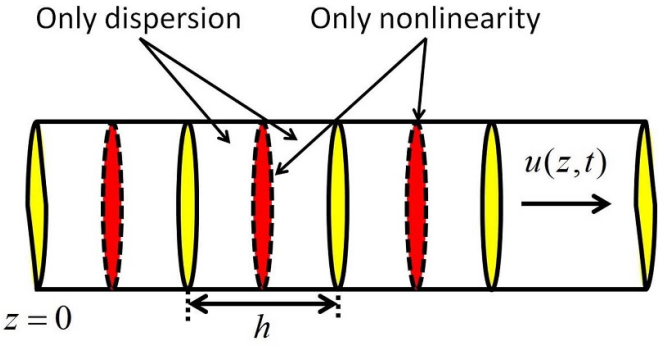

Figure 2. Schematic illustration of the symmetric SSFM.

Example 7 . Show that the symmetric SSFM in (88) is accurate to the third-order in h.

Solution. Using the Baker–Hausdorff formula in [28], we can write:

$$\begin{aligned}\exp\left(\frac{\mu}{2}A\right)\exp(\mu B)\exp\left(\frac{\mu}{2}A\right) &= \left(\exp\left(\frac{\mu}{2}A\right)\exp(\mu B)\right)\exp\left(\frac{\mu}{2}A\right)\\ &= \exp\left(\frac{\mu}{2}A + \mu B + \frac{\mu^2}{4}[A,B] + \frac{\mu^3}{24}\left[\frac{A}{2}-B,[A,B]\right] + \cdots\right)\\ &\quad\times\exp\left(\frac{\mu}{2}A\right).\end{aligned} \tag{89}$$

By the substitution $C = \frac{A}{2} + B + \frac{\mu}{4}[A,B] + \frac{\mu^2}{24}\left[\frac{A}{2}-B,[A,B]\right] + \cdots$, (89) can be rewritten as:

$$\begin{aligned}\exp\left(\frac{\mu}{2}A\right)\exp(\mu B)\exp\left(\frac{\mu}{2}A\right) &= \exp(\mu C)\exp\left(\frac{\mu}{2}A\right)\\ &= \exp\left(\mu C + \frac{\mu}{2}A + \frac{\mu^2}{4}[C,A] + \frac{\mu^3}{24}\left[C-\frac{A}{2},[C,A]\right] + \cdots\right).\end{aligned} \tag{90}$$

Considering the part of (90) corresponds to the term μ^2 and substituting the expression for C, we obtain:

$$\frac{\mu^2}{4}[C,A] = \frac{\mu^2}{4}\left[\left(\frac{A}{2} + B + \frac{\mu}{4}[A,B] + \frac{\mu^2}{24}\left[\frac{A}{2}-B,[A,B]\right] + \cdots\right),A\right]. \tag{91}$$

By expanding (91) following the substitution of $\mu = h, A = \hat{D}$, $B = \hat{N}$ and considering only the dominant error term by ignoring all other terms (note that this assumption is only to isolate the dominant error term for the mathematical simplicity of explanation), we obtain the term $\frac{h^3}{16}\left[\hat{D},\left[\hat{D},\hat{N}\right]\right]$. From this, we can conclude that the symmetric SSFM is accurate to the third-order in h, especially for small values of h.

The symmetric SSFM in (88) has some important advantages over the SSFM implementation in (84), including:

(a) The integral in the nonlinear section (i.e. the middle exponential) is useful to incorporate the z dependence of $\hat{N}$, especially for large step size h. Also, it can be approximated as $\exp\left(h\hat{N}\right)$ for small values of h.
(b) The symmetric SSFM is accurate to the third-order in the step size h compared to the SSFM in (84), which is accurate only to the second-order in h. This fact is evident from examples 6 and 7.

The implementation of the symmetric SSFM is relatively straightforward [1]. The fiber span is divided into a large number of segments of size h. The optical signal is propagated from segment to segment as formulated in (88). More specifically, the optical field envelop $u(z,t)$ is first propagated for a distance $h/2$ with only dispersion effect. At the midplane, the optical field is multiplied by the nonlinear term that represents the nonlinearity effect for the whole segment of length h. Finally, the optical field is propagated the remaining distance of length $h/2$ with only dispersion effect to obtain the optical field $u(z+h,t)$ [1].

5.2. Analytical approximation methods

The SSFM-based numerical solution of the NLSE provides an accurate model for the linear and nonlinear signal propagation in the optical fiber; however, its computational complexity is impractically high [1]. This drawback of the SSFM motivates the research community to search for approximation methods to solve the NLSE analytically. The most commonly used analytical approximations used for the solution of NLSE are the Volterra series-based and the regular perturbation (RP) series-based methods.

5.2.1. Volterra series-based method. The Volterra series-based analysis is a well-established tool used in the nonlinear systems theory and extensively used to solve nonlinear problems in the field of communications engineering [29–35]. The seminal work of Peddanarappagari and Brandt-Pearce in [34] to analytically approximate the solution of NLSE opens the door of further research interests in the approximate solutions of the fiber propagation equation. These research efforts initiate the adaptation of known communication theory results, such as the equalization techniques for nonlinear channels, for use in the nonlinear optical fiber channels [32].

The Volterra series-based approach can be used to approximate the solution of NLSE analytically. In this approach, the relation between the input and the output of the optical fiber can be represented as a series of transfer function kernels, referred to as Volterra series transfer functions (VSTFs) [29–35]. The VSTF provides the relationship between the Fourier transforms of the input to the fiber $X(w)$ and the output at the end of the fiber link $Y(w)$, as [29–35]:

$$Y(w) = H_1(w)X(w) + \sum_{n=2}^{\infty} \int \cdots \int H_n(w_1,\ldots,w_{n-1},w) \times X(w_1)\cdots X(w_{n-1})X(w-w_1-\cdots-w_{n-1}) \times dw_1\cdots dw_{n-1}, \tag{92}$$

where $H_n(w_1,\ldots,w_{n-1},w)$ is the nth order VSTF kernel in the frequency-domain. Due to the symmetries in the silica-based optical fiber, the even-ordered Volterra kernels are zero. The Volterra series is similar to the Taylor series; however, the former captures the memory effect of the nonlinear system, and the latter does not. The VSTF in (92) includes the memory effects caused by the fiber CD and the CD-nonlinearity interplay. More specifically, the VSTF kernel labeled H_1 represents the linear or CD-only response of the fiber channel, and the higher-order kernel terms $H_{2,3,\ldots,n}$ defines the higher-order nonlinear transfer functions, including the CD-nonlinearity interplay. It is important to mention that the VSTF kernels up to third-order provide an acceptable level of accuracy for the typical values of peak launch power and the transmission distance used in the optical fiber communication systems.

For a PDM transmission system, the VSTF kernels up to third-order can be represented as [29–35]:

$$U_x(z,w) = H_1(z,w)U_x(w) + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H_3(z,w_1,w_2,w) \times \left[U_x(w_1)U_x^*(w_2) + U_y(w_1)U_y^*(w_2)\right] \times U_x(w-w_1+w_2)\,dw_1dw_2, \tag{93}$$

where $U_{x/y}(w) \triangleq U_{x/y}(z=0,w)$ represents the Fourier transform of the field envelop $u_{x/y}(z,t)$ and VSTF kernels can be given as:

$$H_1(z,w) = \exp\left(-\frac{\alpha}{2} - \frac{jw^2\beta_2}{2}\right)z, \tag{94}$$

$$H_3(z,w_1,w_2,w) = j\frac{8}{9}\frac{\gamma}{4\pi^2}H_1(z,w) \times \frac{1-\exp\left(-(\alpha+j\beta_2(w-w_1)(w_1-w_2))z\right)}{\alpha+j\beta_2(w-w_1)(w_1-w_2)}. \tag{95}$$

Substituting (94) and (95) in (93) and detaching $H_1(z,w)$ from $H_3(z,w_1,w_2,w)$, we can modify (93) as [29–35]:

$$U_x(z,w) = H_1(z,w)U_x(w) + H_1(z,w)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_3(z,w_1,w_2,w) \times \left[U_x(w_1)U_x^*(w_2) + U_y(w_1)U_y^*(w_2)\right] \times U_x(w-w_1+w_2)\,dw_1dw_2, \tag{96}$$

where K_3 is defined as:

$$K_3(z,w_1,w_2,w) = j\frac{8}{9}\frac{\gamma}{4\pi^2} \times \frac{1-\exp\left(-(\alpha+j\beta_2(w-w_1)(w_1-w_2))z\right)}{\alpha+j\beta_2(w-w_1)(w_1-w_2)}. \quad (97)$$

It is understood from (96) that the pulse propagation in SSMF is modelled by the combination of a linear kernel and a third-order nonlinear kernel. The linear kernel $H_1(z,w)$ models the dispersion and attenuation effect in the optical fiber; whereas, the third-order nonlinear kernel $H_3(z,w_1,w_2,w)$ models the beating of various frequency components present in the input signal. It is important to mention that, for long-haul links, the complexity of the Volterra series-based method increases corresponding to the increase in the number of spans [29–35].

5.2.2. Perturbation theory-based method. The RP theory-based approach has been proposed to analytically approximate the solution of NLSE with a reduced complexity when compared to the Volterra series-based approach. The RP-based solution provides a detailed insight into the complex interaction between the Kerr nonlinearity and the accumulated CD in the optical fiber. In this method, the signal field $u(z,t)$ is expanded in a power series of γ as $u(z,t) = \sum_{k'=0}^{\infty}\gamma^{k'}u_{k'}(z,t)$, where k' is the order of the solution. Then, substituting $u(z,t)$ in (73), we get [36]:

$$\sum_{k'=0}^{\infty}\gamma^{k'}\frac{\partial}{\partial z}u_{k'}(z,t) = -\sum_{k'=0}^{\infty}\gamma^{k'}j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}u_{k'}(z,t) + j\gamma\sum_{m=0}^{\infty}\sum_{l=0}^{\infty}\sum_{n=0}^{\infty}\gamma^{m+l+n}\times u_m(z,t)u_l^*(z,t) \times u_n(z,t)\exp(-\alpha z), \quad (98)$$

where γ represents the nonlinearity coefficient, m, n, l indicate the symbol indices, α denotes the attenuation coefficient, and z stands for the transmission distance. We obtain a set of linear differential equations by equating the terms in (98) that have equal powers of γ on both sides of the equal sign. Accordingly, the differential equation for the k'th order solution is given as [36]:

$$\frac{\partial}{\partial z}u_{k'}(z,t) = -j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}u_{k'}(z,t) + j\sum\sum_{m+l+n=k'-1}\sum u_m(z,t)u_l^*(z,t)u_n(z,t) \times \exp(-\alpha z). \quad (99)$$

5.2.3. Zeroth-order (or linear) solution. By substituting $k'=0$ in (99), the differential equation for the zeroth-order (or linear) solution can be written as [37, 38]:

$$\frac{\partial}{\partial z}u_0(z,t) = -j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}u_0(z,t). \quad (100)$$

By solving (100), the zeroth-order solution at $z=L$ can be written as [37, 38]:

$$u_0(L,t) = u(0,t)\otimes h_L(t), \quad (101)$$

where $\otimes$ represents the convolution, $h_L(t) = \mathcal{F}^{-1}\{\exp(-j\frac{w^2\beta_2 L}{2})\} = \frac{1}{\sqrt{-2\pi j\beta_2 z}}\exp\left(\frac{-jt^2}{2\beta_2 z}\right)$ with w as the angular frequency and $\mathcal{F}^{-1}\{.\}$ as the inverse Fourier transform function.

5.2.4. First-order solution. The differential equation for the first-order solution is obtained by substituting $k'=1$ in (99) and can be represented as [37, 38]:

$$\frac{\partial}{\partial z}u_1(z,t) = -j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}u_1(z,t) + j|u_0(z,t)|^2u_0(z,t)\exp(-\alpha z). \quad (102)$$

The first-order distortion field at a transmission distance $z=L$ is obtained by solving (102) and assuming an ideal dispersion compensation at $z=L$, we get [37, 38]:

$$u_1(L,t) = j\gamma\int_0^L\left(h_z(t)\otimes[|u_0(z,t)|^2u_0(z,t)]\right)\exp(-\alpha z)dz. \quad (103)$$

The frequency-domain distortion field at $z=L$ is obtained by taking the Fourier transform of (103) and the result is given as [37, 38]:

$$U_1(L,w) = j\gamma\int_0^L F(z,w)\exp(-j\frac{w^2\beta_2 z}{2})\exp(-\alpha z)dz, \quad (104)$$

where $F(z,w)$ is written as [37, 38]:

$$F(z,w) = \int_{-\infty}^{\infty}|u_0(z,t)|^2u_0(z,t)\exp(-jwt)\,dt. \quad (105)$$

The signal input to the optical fiber can be represented as [37, 38]:

$$u(z=0,t) = \sqrt{P_0}\sum a_k g(z=0,t-kT), \quad (106)$$

where P_0 is the peak launch power, a_k is the data symbol of the kth pulse, $g(z,t)$ is the pulse temporal waveform at transmission distance z, and T is the symbol duration. Therefore, the product $|u_0(z,t)|^2u_0(z,t)$ in (105) can be represented as [37, 38]:

$$|u_0(z,t)|^2u_0(z,t) = P_0^{3/2}\sum_m\sum_l\sum_n a_m g(z,t-mT)a_l^* \times g^*(z,t-lT)a_n g(z,t-nT) = P_0^{3/2}\sum_m\sum_l\sum_n a_m g(z)a_l^*g^*(z)a_n g(z). \quad (107)$$

Substituting (107) in (105), we obtain [36]:

$$\begin{aligned}F(z,w) &= P_0^{3/2}\int_{-\infty}^{\infty}\left(\sum_m\sum_l\sum_n a_m g(z) a_l^* g^*(z) a_n g(z)\right)\\ &\quad\times\exp(-jwt)\,dt\\ &= P_0^{3/2}\sum_m\sum_l\sum_n a_m a_l^* a_n (G_m(z,w)\otimes G_l^*(z,-w)\\ &\quad\otimes G_n(z,w))\\ &= P_0^{3/2}\sum_m\sum_l\sum_n a_m a_l^* a_n \exp\left(j\frac{w^2\beta_2 z}{2}\right)\\ &\quad\times\exp(-jw(T_m-T_l+T_n))\\ &\quad\times\int\int G(0,w_1+w)G^*(0,w_1+w-w_2)\\ &\quad\times G(0,w-w_2)\exp(j\beta_2 z w_1 w_2)\\ &\quad\times\exp(-j(w_1(T_m-T_l)+w_2(T_l-T_n)))\,dw_1dw_2,\end{aligned}\tag{108}$$

where $G(z,w)=G(0,w)\exp\left(j\frac{w^2\beta_2 z}{2}\right)$ with $G(0,w)$ is the Fourier transform of $g(0,t)$.

Substituting (108) in (104), we get first-order perturbation kernel term in frequency-domain as [37, 38]:

$$\begin{aligned}U_1(L,w) &= j\gamma P_0^{3/2}\sum_m\sum_l\sum_n a_m a_l^* a_n\exp(-jw(T_m-T_l+T_n))\\ &\quad\times\int_0^L\exp(-\alpha z)\left(\int\int G(0,w_1+w)G^*\right.\\ &\quad\times(0,w_1+w-w_2)G(0,w-w_2)\\ &\quad\times\exp(-j(w_1(T_m-T_l)+w_2(T_l-T_n)))\\ &\quad\left.\times\exp(jw_1w_2\beta_2 z)\,dw_1dw_2\right)dz.\end{aligned}\tag{109}$$

Next, by taking the inverse Fourier transform of (109), the time-domain first-order kernel term can be represented as [37, 38]:

$$\begin{aligned}&u_1(L,(t+T_m-T_l+T_n))\\ &= j\gamma P_0^{3/2}\sum_m\sum_l\sum_n a_m a_l^* a_n\int_0^L\exp(-\alpha z)\\ &\quad\times\left(\int\int\int G(0,w_1+w)G^*(0,w_1+w-w_2)\right.\\ &\quad\times G(0,w-w_2)\exp(-j(w_1(T_m-T_l)+w_2(T_l-T_n)))\\ &\quad\left.\times\exp(jw_1w_2\beta_2 z)\exp(jwt)\,dw_1dw_2dw\right)dz.\end{aligned}\tag{110}$$

For Gaussian pulse shape assumption, i.e. $G(0,w)=\sqrt{2\pi\tau^2}\exp(-\frac{w^2\tau^2}{2})$, with τ as the pulse width, the triplet product in (110) can be calculated as [37, 38]:

$$\begin{aligned}&G(0,w_1+w)G^*(0,w_1+w-w_2)G(0,w-w_2)\\ &= \left(\sqrt{2\pi\tau^2}\right)^3\exp\left(-\frac{3\tau^2w^2}{2}\right)\\ &\quad\times\exp(-\tau^2[w_1^2+w_2^2+2(w_1-w_2)w-w_1w_2]).\end{aligned}\tag{111}$$

It is important to mention that with the Gaussian shape assumption for the input pulse shape, the first-order nonlinearity coefficients can be calculated using analytic expressions, which involve the exponential integral function [39]. This will explain in detail in section 7.3.

Substituting (111) in (110) and integrating w.r.t w, we get [37, 38]:

$$\begin{aligned}u_1(L,(t+T_m-T_l+T_n)) &= j\gamma P_0^{3/2}\frac{2\pi\tau^2}{\sqrt{3}}\exp\left(-\frac{t^2}{6\tau^2}\right)\\ &\quad\times\sum_m\sum_l\sum_n a_m a_l^* a_n\\ &\quad\times\int_0^L\int\int\exp(-\alpha z)\\ &\quad\times\exp\left(-\frac{1}{3}\tau^2(w_1^2+w_2^2+w_1w_2)\right)\\ &\quad\times\exp\left(-j\left(\frac{2}{3}(w_1-w_2)t+w_1(T_m-T_l)\right.\right.\\ &\quad\left.\left.+w_2(T_l-T_n)-w_1w_2\beta_2 z\right)\right.\\ &\quad\left.\times dw_1dw_2dz\right).\end{aligned}\tag{112}$$

Next, integrating (112) and carrying out some algebraic simplifications, we obtain the first-order distortion field (or first-order ghost pulse) as [37, 38]:

$$\begin{aligned}u_1(L,t+kT) &= j\gamma P_0^{3/2}\sum_m\sum_l\sum_n a_m a_l^* a_n\exp\left(-\frac{t^2}{6\tau^2}\right)\\ &\quad\times\int_0^L\frac{\exp(-\alpha z)}{\sqrt{1+2j\beta_2 z/\tau^2+3(\beta_2 z/\tau^2)^2}}\\ &\quad\times\exp\left\{\begin{array}{l}-\frac{3\left[\frac{2}{3}t+(m-l)T\right]\left[\frac{2}{3}t+(n-l)T\right]}{\tau^2(1+3j\beta_2 z/\tau^2)}\\ -\frac{(n-m)^2T^2}{\tau^2[1+2j\beta_2 z/\tau^2+3(\beta_2 z/\tau^2)^2]}\end{array}\right\}dz,\end{aligned}\tag{113}$$

where T_m, T_l, and T_n can be represented as mT, lT, and nT, respectively, $k=m+n-l$, m,n,l are the symbol indices, $a_{m/l/n}$ is the symbol complex amplitude, and τ is the pulse width.

In the PDM optical system, the input signal field to the fiber is a column vector $\mathbf{u}(z,t)=[u_x(z,t)\ u_y(z,t)]^\dagger$, where x, y are the polarization tributaries, and $\dagger$ is the transpose. The vector field propagation in the optical fiber is governed by the Manakov equation, which is given as [40]:

$$\frac{\partial}{\partial z}\mathbf{u}+j\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}\mathbf{u}=j\frac{8}{9}\gamma(\mathbf{u}^{*\dagger}\mathbf{u}\mathbf{I})\mathbf{u}\exp(-\alpha z),\tag{114}$$

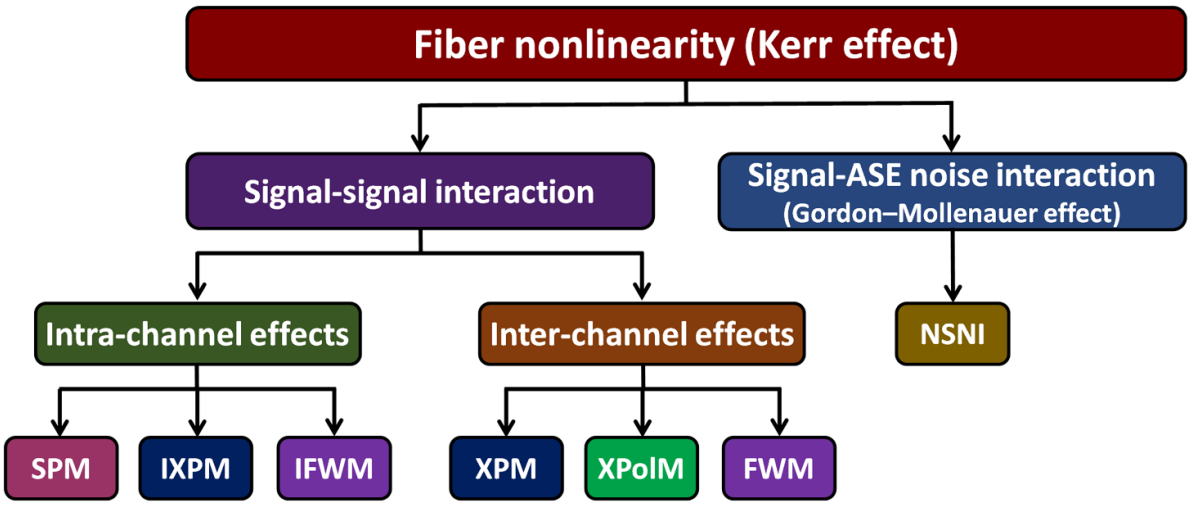

Figure 3. Fiber nonlinearity effects. SPM: self-phase modulation, (I)XPM: (intra-channel) cross-phase modulation, XPolM: cross-polarization modulation, (I)FWM: (intra-channel) four-wave mixing, ASE: amplified spontaneous emission, and NSNI: nonlinear signal-noise interaction.

where **I** is the identity matrix. It is important to note that (114) is the same as the Manakov equation given in (74), in which (114) is written in a more compact form by omitting the space and time variables z, t for the sake of simplicity. By solving (114), the zeroth- and first-order solutions can be written as [40]:

$$u_{0,x/y}(L,t) = u_{x/y}(0,t) \otimes h_L(t), \tag{115}$$

and

$$u_{1,x/y}(L,t) = j\gamma \int_0^L \left(h_z(t) \otimes [\left|u_{0,x/y}(z,t)\right|^2 u_{0,x/y}(z,t)] \right) \times \exp(-\alpha z) dz. \tag{116}$$

Following the steps from (104) to (113), we get the first-order ghost pulse for the PDM transmission system as [37, 38]:

$$u_{1,x/y}(L,t+kT) = j\frac{8}{9}\gamma P_0^{3/2} \sum_m \sum_l \sum_n [a_{m,x/y} a^*_{l,x/y} a_{n,x/y} + a_{m,y/x} a^*_{l,y/x} a_{n,x/y}] \times \exp(-\frac{t^2}{6\tau^2}) \int_0^L \frac{\exp(-\alpha z)}{\sqrt{1+2j\beta_2 z/\tau^2 + 3(\beta_2 z/\tau^2)^2}} \times \exp\left\{ \begin{array}{l} -\frac{3\left[\frac{2}{3}t+(m-l)T\right]\left[\frac{2}{3}t+(n-l)T\right]}{\tau^2(1+3j\beta_2 z/\tau^2)} \\ -\frac{(n-m)^2 T^2}{\tau^2\left[1+2j\beta_2 z/\tau^2+3(\beta_2 z/\tau^2)^2\right]} \end{array} \right\} dz. \tag{117}$$

6. Kerr-induced fiber nonlinearity effects

As stated in section 3, the nonlinearity effects in the optical fiber are due to the Kerr effect, which arises from the dependence of the optical fiber refractive index on the transmit signal power [1]. Figure 3 illustrates different types of Kerr nonlinearity effects in the optical fiber. The nonlinearity effects can be divided into two types: signal-signal and signal-amplified spontaneous emission (ASE) noise nonlinear interaction [7]. In single-channel systems, the intra-channel signal-signal nonlinear interactions can be categorized into three types: (a) SPM, (b) IXPM, and (c) IFWM. The SPM results in a phase modulation induced by the refractive index, which increases with the increase in the input power level. This leads to a frequency chirping effect [7, 41–43], which interacts with the CD and causes the spectral broadening of the optical pulse. The IXPM is the result of the refractive index change proportional to the intensity of the neighbouring pulses in the same channel [7, 41–43]. The IXPM yields a timing jitter between the co-propagating pulses, which leads to performance degradation. The IFWM is caused by the nonlinear interaction between two or more pulses of the same channel [7, 41–43]. That generates echo or ghost pulses in the time domain, and thus, results in interference between the signal pulses of the same channel. It is important to mention that the SPM, IXPM, and IFWM can be compensated well using digital NLC techniques [7, 41–43].

In WDM systems, the inter-channel signal-signal nonlinear effects can also be classified into three types: (a) XPM, (b) XPolM, and (c) FWM. The XPM effect is due to the refractive index change proportional to the intensity of the pulse in the co-propagating channel. The XPM induces frequency chirping and pulses overlapping between channels [7, 41–43]. The XPM effect consists of two parts, the coherent and incoherent XPM, as given in (79). The coherent XPM is the same as the usual XPM effect, which involves the interaction of two WDM channels. The incoherent XPM results in polarization cross-talk, and the cross-talk coefficient is determined by the polarization cross-product $u^*_{y/x,m} u_{x/y,m}$ of the interfering channel [7]. The XPolM occurs in PDM-WDM transmission systems when the SOP of a transmitted channel depends on the SOP of other co-propagating channels through Kerr effect. That is

due to the random propagation of SOP inside the optical fiber caused by PMD [7, 41–43] and causes channel cross-talk for PDM-WDM systems [7, 41–43]. The FWM in the multichannel system is generated by the nonlinear interaction between two or more co-propagating wavelength channels [7, 41–43]. The FWM results in significant performance degradation due to cross-talk among different wavelength channels [7, 41–43].

Another nonlinearity effect is the Kerr-induced signal-ASE noise interaction, referred to as the Gordon–Mollenauer effect [44, 45], which leads to the generation of the nonlinear signal-noise interaction (NSNI) [44, 45]. The modeling of the impact of NSNI is important because it can contribute to answering the questions regarding the fundamental limits of the performance improvement provided by the NLC techniques [44, 45].

7. Fiber nonlinearity compensation techniques

The impairments due to the Kerr effect degrade the transmission performance of optical systems using higher-order modulation formats. Therefore, NLC is a hot research topic to increase fiber capacity. Several digital and optical techniques have been reported in the literature to combat the effects of fiber nonlinearity. Such techniques are applied either at the transmitter side or in the optical link, or at the receiver side. The introduction of coherent detection in optical communications initiates the use of DSP algorithms to compensate for fiber impairments. Usually, DSP algorithms are implemented either at the transmitter side or at the receiver side, or a combination of both like split nonlinearity compensation techniques [7]. Implementing DSP algorithms in the optical fiber link requires optical-to-electrical/electrical-to-optical conversions at the locations where the DSP modules are deployed. That increases the signal latency in the optical fiber link and the overall cost of the link.

Digital NLC techniques represent a key technology and a cost-effective approach to increase the data rate, being adopted for the next generation WDM optical transmission systems [7]. In this section, we provide a concise review of the most popular NLC techniques in the literature.

7.1. Digital back-propagation

The digital back-propagation (DBP) is based on the SSFM, which represents an effective numerical technique to solve the signal propagation equation. The idea of the DBP technique is to digitally model a fictitious fiber with exactly opposite characteristics when compared to the real fiber used for the transmission [7]. In this technique, the optical fiber is first divided into small segments, and the signal propagation in each segment is modelled as a concatenation of linear and nonlinear operations [7]. There are different DBP implementations available depending on the order of the linear and nonlinear sections in a given fiber segment [7]. The most commonly used approach is the application of linear compensation first because nonlinear effects are more important at high input powers, which is the case at the end of the fictitious fiber [7]. The receiver side DBP implementation is shown in figure 4, where N_s is the number of steps.

Using SSFM, the output of the CD (or linear) compensation section is given by [7]:

$$U^{\mathrm{CD}}_{x/y}(z,w) = U_{x/y}(z,w)\exp(-jh(\frac{\alpha}{2}+\frac{\beta_2}{2}\omega^2)), \qquad (118)$$

where h is the length of each step, w is the frequency variable and z is the transmission distance. Following that, the nonlinear section is carried out in time-domain as [7]:

$$u'_{x/y}(z,t) = u^{\mathrm{CD}}_{x/y}(z,t)\exp(-j\varphi\gamma' h(|u^{\mathrm{CD}}_x|^2+|u^{\mathrm{CD}}_y|^2)), \quad (119)$$

where $0<\varphi<1$ is a real-valued optimization parameter.

Single-channel DBP (SC-DBP) has been more widely researched because it is typically considered to be more realistic with the current hardware limitations [7]. Since a single-wavelength channel is back-propagated, SC-DBP only compensates for intra-channel nonlinearity, e.g. SPM. In WDM superchannel systems, one way of overcoming the inter-channel nonlinear distortions like XPM and FWM induced by the co-propagating subchannels is to use a multi channel-DBP (MC-DBP). The MC-DBP back-propagates the entire WDM channel [7]. However, the implementation of the MC-DBP is limited to point-to-point links, and its computational complexity is considered impractical due to the need for massively parallel processing computer systems to implement [7]. There are a few reduced complexity DBP implementations are available, such as weighted DBP [7] and correlated DBP [7]. However, their computational complexity is still considered high. Even though the DBP implementation complexity is high, it is considered a benchmark to measure the effectiveness of other NLC techniques.

7.2. Volterra series-based nonlinear equalizer

The VSTF can be effectively used to model the fiber nonlinearity effects [7]. In this technique, after modelling the optical channel as a series of VSTFs, the pth order theory proposed in [33] is adopted to design the inverse VSTF (IVSTF) kernels. These IVSTF kernels can be used to compensate for the fiber Kerr nonlinearity and its interplay with CD. One of the features of the Volterra series-based nonlinear equalizer (VNLE) is that the compensation operation can be performed in parallel [7]. That reduces the computational complexity when compared to DBP [7]. The principle of the parallel implementation of VNLE is depicted in figure 5, where N_{spans} corresponds to the number of fiber spans.

For each polarization, the compensation operation can be divided into two parts; one is linear, and the other is nonlinear. The linear part consists of CD compensation, and the nonlinear part compensates for the nonlinearity effect. In this technique, the nonlinearity compensation for each span can be carried out in parallel. We obtain the output compensated signal by combining the output of the linear section and the output of each nonlinear stage, as shown

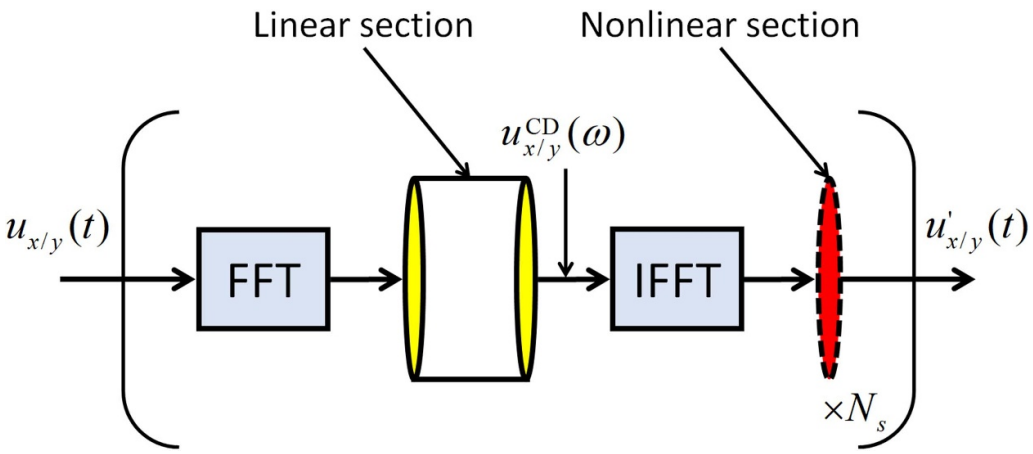

Figure 4. Principle of digital back-propagation (DBP) implementation. (I)FFT: (inverse) fast Fourier transform.

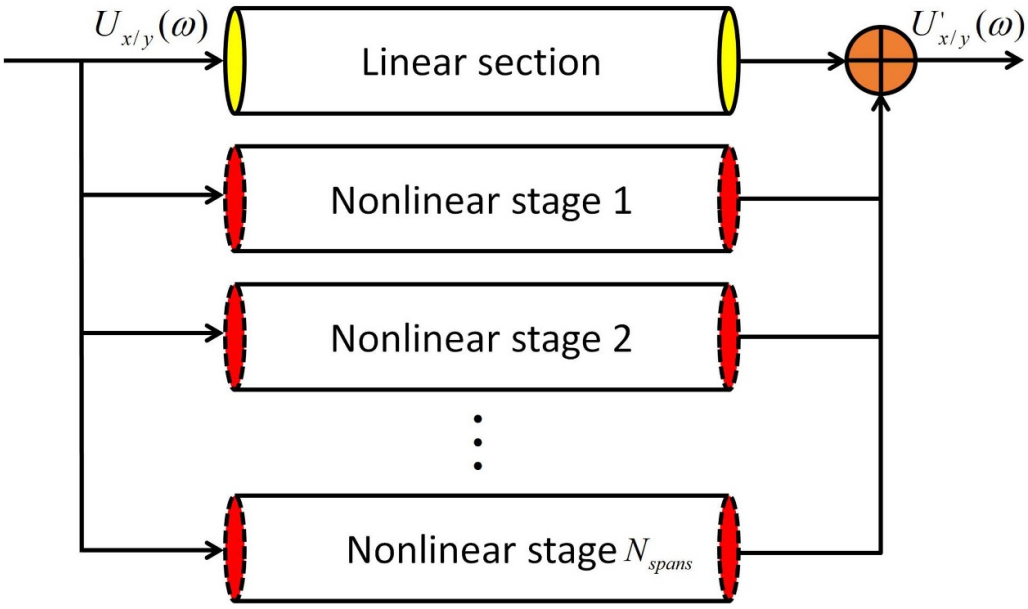

Figure 5. Parallel implementation of VNLE.

in figure 5. In the literature, there are different variants of VNLE have been proposed, such as modified VNLE (MVNLE) and weighted Volterra series nonlinear equalizer (W-VSNE), to improve the compensation performance and reduce the implementation complexity [7]. It is worth noting that the VNLE, MVNLE, and W-VSNE are based on the third-order Volterra series approximation. A fifth-order VNLE has also been proposed in the literature to compensate for the Kerr nonlinearity effect; however, it involves much higher implementation complexity when compared to its third-order counterpart [7].

7.3. Perturbation theory-based NLC

The first-order perturbation theory-based NLC (PB-NLC) technique adopts some simplifying assumptions, including [38, 40, 46–49]:

- The full electronic compensation of CD at the receiver.
- The Gaussian shape assumption for input pulses.

According to the first-order theory, three input Gaussian pulses $\sqrt{P_0}a_{m/l/n,x/y}\exp(-(t-T_{m/n/l})^2/2\tau^2)$, at time indices T_m, T_l, T_n interact nonlinearly and generate a ghost pulse, as shown in (117). Figure 6 shows the schematic of the triplet interaction to generate the first-order field. Without loss of generality, the nonlinear distortion field induced at index $k=0$, i.e. $l=m+n$ is calculated at $t=0$ by considering the symbol rate operation [38, 40, 46–49]. Accordingly, (117) can be further simplified as [37, 38, 40, 46–49]:

$$u_{1,x/y}(L,t) = j\frac{8}{9}\gamma P_0^{3/2}\sum_m\sum_n\left[a_{m,x/y}a^*_{m+n,x/y}a_{n,x/y} + a_{m,y/x}a^*_{m+n,y/x}a_{n,x/y}\right]\mathbf{C}_{m,n}, \tag{120}$$

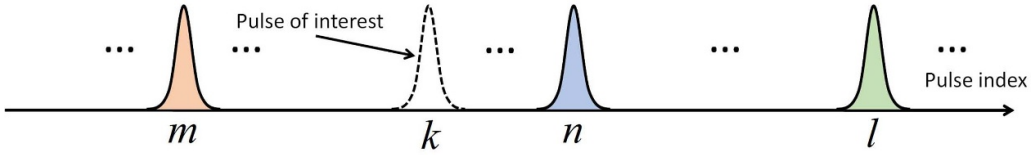

Figure 6. Triplet pulses in the PB-NLC technique.

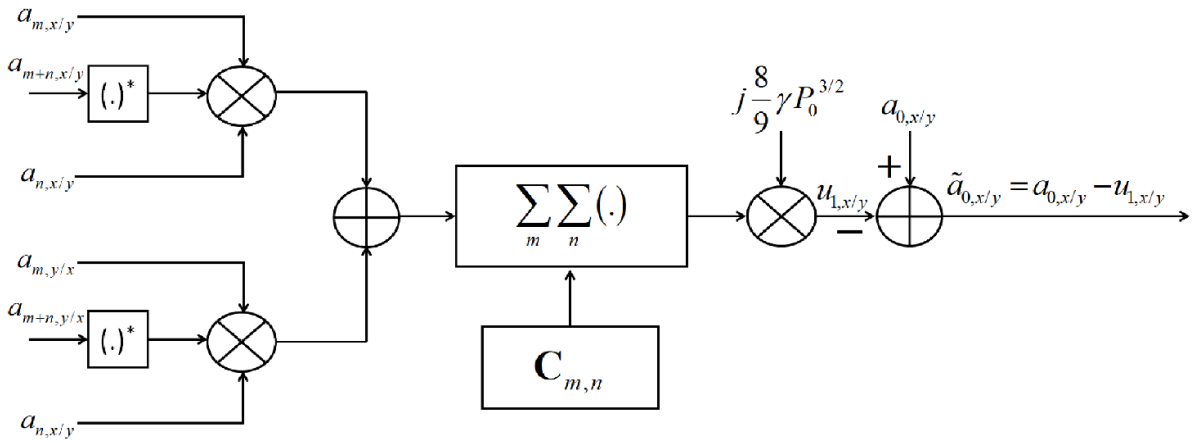

Figure 7. Principle of operation of the PB-NLC technique.

where $*$ represents the complex conjugate operation and $\mathbf{C}_{m,n}$ is the first-order perturbation coefficient matrix, which is given as [37, 38, 40, 46–49]:

$$\mathbf{C}_{m,n} = \int_0^L \frac{\exp(-\alpha z)}{\sqrt{1+2j\beta_2 z/\tau^2 + 3(\beta_2 z/\tau^2)^2}} \times \exp\left(-3\frac{mnT^2}{\tau^2(1+3j\beta_2 z/\tau^2)} - \frac{(m-n)^2 T^2}{\tau^2[1+2j\beta_2 z/\tau^2 + 3(\beta_2 z/\tau^2)^2]}\right) dz. \quad (121)$$

In a typical dispersion uncompensated system, the pulse spreading due to CD is much higher than the symbol duration, i.e. $\beta_2 z \gg \tau^2$ [38]. Using this assumption and following the simplification steps in [38], the first-order perturbation coefficient matrix $\mathbf{C}_{m,n}$ can be written as [37, 38, 40, 46–49]:

$$\mathbf{C}_{m,n} = \begin{cases} \frac{\tau^2}{\sqrt{3}|\beta_2|} \int_0^L dz \frac{1}{\sqrt{\tau^4/(3\beta_2^2)+z^2}}, & m=n=0 \\ \frac{\tau^2}{\sqrt{3}|\beta_2|} \frac{1}{2} E_1\left(\frac{(n-m)^2 T^2 \tau^2}{3|\beta_2|^2 L^2}\right), & m \text{ or } n = 0 \\ \frac{\tau^2}{\sqrt{3}|\beta_2|} E_1\left(-j\frac{mnT^2}{\beta_2 L}\right), & m \neq n \neq 0, \end{cases} \quad (122)$$

where $E_1(x) = \int_x^\infty \frac{e^{-t}}{t} dt$ is the exponential integral function [38].

In the PB-NLC technique, the perturbation coefficients are calculated beforehand and saved in a look-up table. The predistortion technique is implemented by calculating the first-order distortion field $u_{1,x/y}$ using (120) and subtract it from the symbol of interest $a_{0,x/y}$ to generate $\tilde{a}_{0,x/y}$, as shown in figure 7.

7.4. Phase conjugation-based NLC

A few years ago, a digital phase conjugation (DPC) based technique, referred to as the phase-conjugated twin wave (PCTW), was proposed for the mitigation of the first-order nonlinear distortions in PDM optical transmission systems. However, the performance improvement of the PCTW technique comes at the expense of 50% spectral efficiency reduction [50, 51]. Figure 8 shows the PCTW scheme applied for a coherent optical orthogonal frequency-division multiplexing (CO-OFDM) system. The data and its phase-conjugate are transmitted on the same subcarrier frequency of the x- and y-polarizations. At the receiver side, the coherent superposition of the phase-conjugate pairs is carried out to cancel the first-order nonlinear distortion, as illustrated in figure 8.

A time-domain implementation of the generalized PCTW was reported in [52], and this method is referred to as conjugate data repetition (CDR). In the CDR technique, each time-domain signal datum is followed by its conjugate pairs, and the received signals in the adjacent time slots are coherently superimposed at the receiver. Since the nonlinearity interference coefficients change slowly in a highly dispersive channel, the nonlinear distortions generated by conjugate repetition data can be self-canceled by superimposing [52]. In [52], a theoretical explanation based on time-domain perturbation analysis has been provided for the nonlinear distortion cancellation with the CDR technique. It is also shown that the PCTW and the CDR techniques have similar performance,

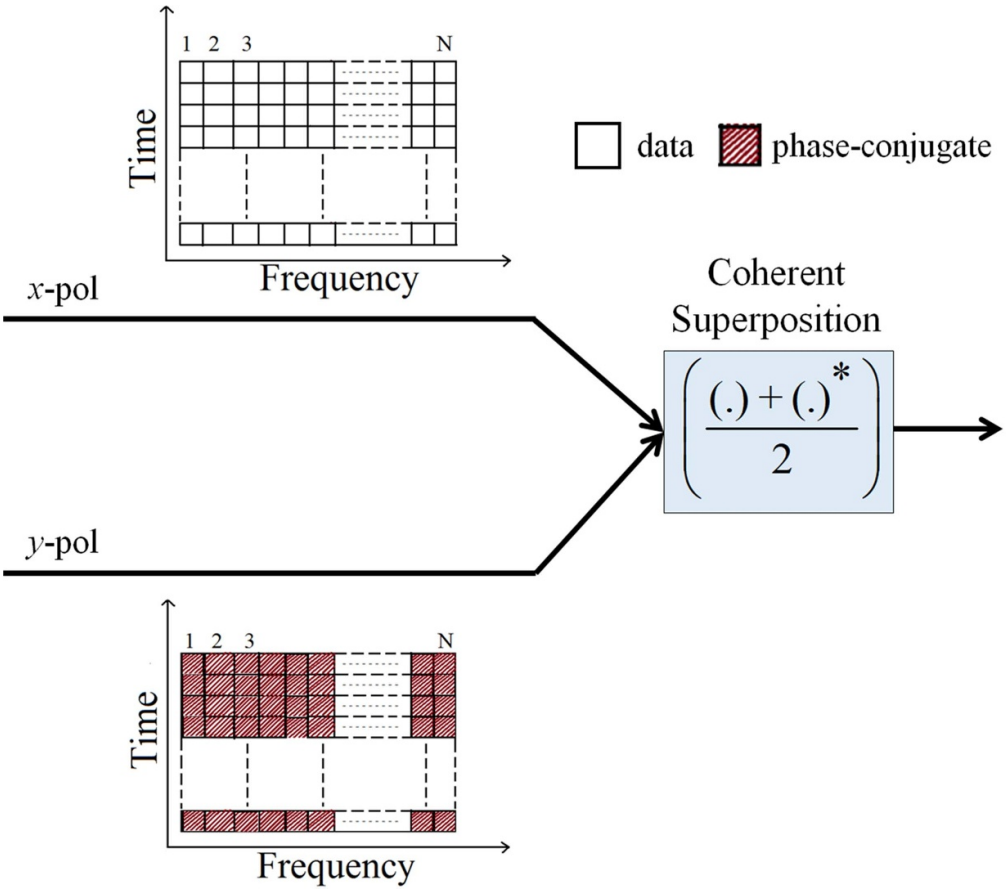

Figure 8. Principle of the PCTW technique. N is the total number of OFDM subcarriers.

with a limitation of halving the overall capacity of the coherent optical communication system. The PCTW technique for the CO-OFDM system utilizing the Hermitian symmetry has been proposed in [53] at the expense of 50% spectral efficiency loss.

7.4.1. DPC techniques with improved spectral efficiency. In [54], a spectrally efficient DPC technique for nonlinearity compensation in the CO-OFDM system has been proposed. This technique is based on the transmission of the phase-conjugated pilots (PCPs) and the coherent superposition at the receiver. In this scheme, a portion of the OFDM subcarriers is transmitted as PCPs of the other subcarriers. These phase conjugate pairs are used at the receiver to estimate and compensate for the nonlinear distortions introduced by the channel. In this technique, the spectral redundancy can be adjusted up to 50% according to the targeted performance gain. That can be achieved through the proper selection of the number of PCPs in each OFDM band. In [55], a novel technique is proposed based on the joint processing of two pairs of PCTWs, referred to as dual-PCTW, to avoid the loss of spectral efficiency associated with the use of PCTWs.

In [56], a frequency-domain coding technique, termed as phase-conjugated subcarrier coding (PCSC), combined with electronic dispersion pre-compensation (pre-EDC), has been demonstrated for nonlinearity mitigation in CO-OFDM system. This technique extends the idea of the dual-PCTW concept to process the neighboring OFDM subcarriers jointly. The benefit of this nonlinearity mitigation technique comes from the fact that the nonlinear distortions on neighboring OFDM subcarriers are highly correlated [56]. The PCSC scheme cancels the first-order distortion field through coherent superposition without losing spectral efficiency. In [57], two DPC approaches have been proposed to overcome the spectral efficiency issue associated with the CDR and PCTW techniques and the pre-EDC requirement of the PCSC technique. These techniques are referred to as the modified-16-quadrature amplitude modulated (QAM)-CDR (MOD-16-QAM-CDR) and MOD-16-QAM-PCTW techniques.

Table 1 demonstrates a list of the state-of-the-art fiber NLC techniques available in the literature to deal with the adverse effect of the Kerr nonlinearity effect. It also shows the type, location, system, and fiber nonlinearity they compensate for, along with the corresponding list of references.

8. Numerical simulations and discussions

In this section, we carry out numerical simulations of the selected nonlinearity compensation techniques discussed in section 7. The selected techniques are: (a) DBP, (b) VNLE, (c) PB-NLC, (d) PCSC, and (e) MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques. In the numerical simulation, we

Table 1. State-of-the-art fiber NLC techniques.

Technique	Type	Location	Nonlin. compensated	System	References
VNLE	Digital	Tx/Rx	Intra-subcarrier	Nyquist/OFDM	[31–35]
PB-NLC	Digital	Tx/Rx	Intra-subcarrier/XPM	Nyquist/OFDM	[36–40, 44, 58]
O/DPC	Optical/Digital	Link/Rx	Nonlinear phase	Nyquist/OFDM	[50–57, 59–61]
SC-DBP	Digital	Tx/Rx	Intra-subcarrier	Nyquist/OFDM	[62–67]
MC-DBP	Digital	Tx/Rx	Intra-/inter subcarrier	Nyquist/OFDM	[68–72]

Table 2. Simulation parameters [38, 48, 49, 57].

Parameter	Value
Root-raised cosine filter roll-off factor	0.1
Fiber span length	80 km
α	0.2 dB km^{-1}
β_2	-20.47 ps^2 km^{-1}
γ	1.2 (1/W) km^{-1}
Noise figure of EDFA	5.5 dB

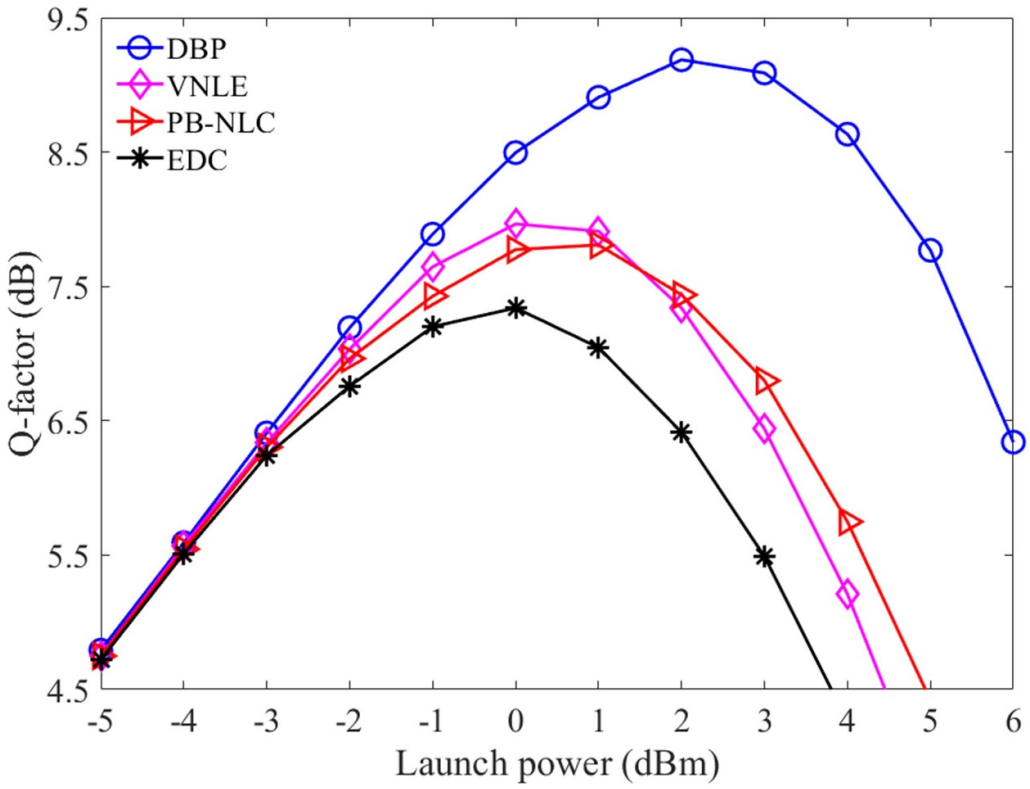

Figure 9. Q-factor as a function of the launch power for DBP, VNLC, PB-NLC, and EDC techniques for a quasi-Nyquist single-carrier system at a transmission distance of 2800 km.

consider both quasi-Nyquist single-carrier and CO-OFDM systems to evaluate the performances of the selected techniques. Specifically, the single-carrier system is used to evaluate the performance of the DBP, VNLE, and PB-NLC techniques. On the other hand, the CO-OFDM system is used to evaluate the performance of the PCSC and MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques. For more details on the system setup for the quasi-Nyquist single-carrier system, please refer to [49]. Similarly, refer to [57] for details on the CO-OFDM system setup. The general parameters used for the numerical simulation are listed in table 2.

Figure 9 shows the Q-factor plotted as a function of the launch power for the DBP, VNLE, PB-NLC, and EDC techniques for the quasi-Nyquist single-carrier system at a transmission distance of 2800 km. It is observed that the DBP technique improves the Q-factor performance by $\sim$1.8 dB when compared to the EDC technique. Also, it can be seen from figure 9 that the VNLE and the PB-NLC techniques show similar Q-factor performances [48, 49].

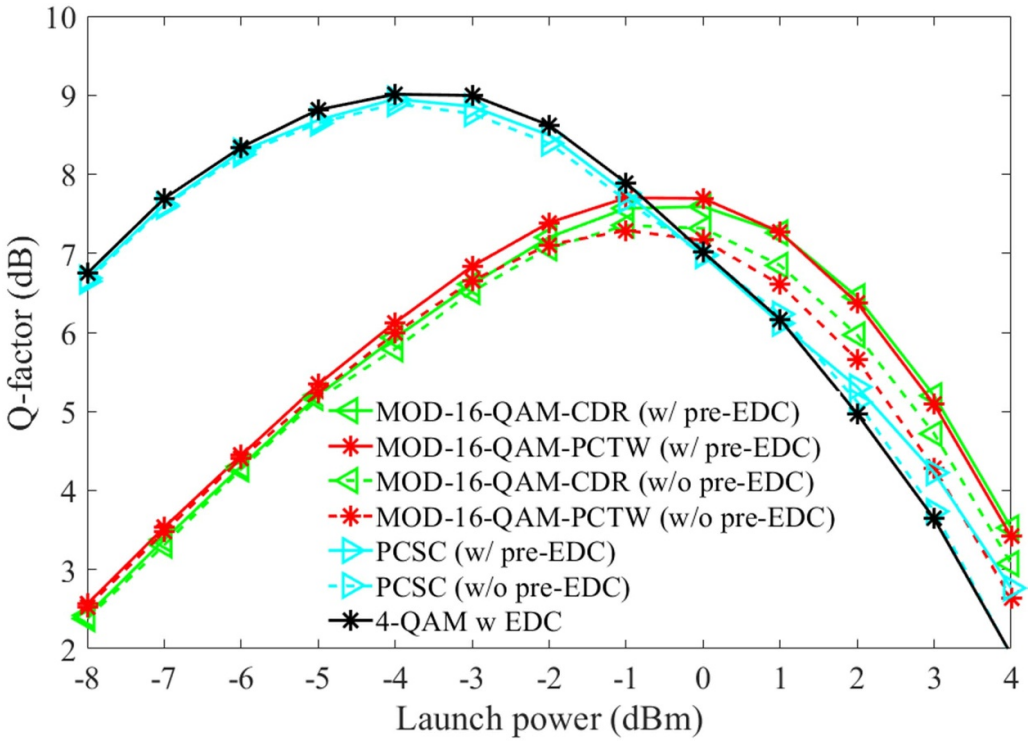

Figure 10. Q-factor as a function of the launch power for MOD-16-QAM-CDR/MOD-16-QAM-PCTW, PCSC, and EDC techniques for a CO-OFDM system at a transmission distance of 2800 km.

Figure 10 shows the Q-factor performances of the MOD-16-QAM-CDR/MOD-16-QAM-PCTW and PCSC techniques with and without considering the pre-EDC for a CO-OFDM system at a transmission distance of 2800 km. The result of the 4-QAM with EDC having the same spectral efficiency is also included for the comparison. Results indicate that the MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques improve the Q-factor performance when compared to the EDC and PCSC techniques at the nonlinear regime of the optical fiber [48, 57]. The results also indicate that the MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques show considerable Q-factor improvement in the nonlinear regime without using pre-EDC when compared to the PCSC technique, which shows negligible performance improvement in this scenario [57].

Table 3 shows the complexity expressions to compute the number of multiplications/symbol for the considered NLC techniques. In table 3, the parameters N_{span} is the number of fiber spans, N_{steps} is the number of steps/span, N_{FFT} is the fast Fourier transform size, P_n is the number of significant perturbation coefficients, and M is the modulation cardinality.

Figure 11 shows the plot of the number of real-valued multiplications/symbol for DBP, VNLE, PB-NLC, and EDC techniques as a function of the number of fiber spans N_{span} for a quasi-Nyquist single-carrier system. In figure 11, $N_{\text{FFT}} = 4096$, and $N_{\text{steps}} = 1$. It is observed that the complexity of the VNLE is slightly higher than the complexity of the PB-NLC technique [7, 48].

Figure 12 shows the number of real-valued multiplications/symbol for MOD-16-QAM-CDR/MOD-16-QAM-PCTW, PCSC, and EDC techniques as a function of the number of fiber spans N_{span} for a CO-OFDM system. It is evident from figure 12 that the performance benefit of the MOD-16-QAM-CDR/MOD-16-QAM-PCTW techniques comes with an increase in the computational complexity [48, 57]. It is important to note that the MOD-16-QAM-CDR/MOD-16-QAM-PCTW and the PCSC techniques are comparatively low complexity NLC techniques, and their complexity is independent of the number of fiber spans N_{span} [48, 57].

It is important to mention that the NLC techniques implemented in this section do not consider the effects of the PMD and the laser phase noise to clearly understand the ability of the selected NLC techniques to deal with the fiber nonlinearity effect. However, in reality, the performance of the NLC techniques may be impacted by such impairments. Several works in the literature carried out the investigation on the impact of the PMD on the NLC techniques. In [73], an investigation of the impact of PMD on the performance of

Table 3. Complexity expressions [48, 49, 57].

Algorithm	Complexity expression
DBP	$N_{\text{span}}N_{\text{steps}}(8\log_2(N_{\text{FFT}})+21)$
VNLE	$N_{\text{span}}N_{\text{steps}}(4\log_2(N_{\text{FFT}})+8.5)$
PB-NLC	$2(4P_n+3)$
MOD-16-QAM-CDR/MOD-16-QAM-PCTW	$8\log_2(N_{\text{FFT}})+4M+9$
PCSC	$8(\log_2(N_{\text{FFT}})+1)+1$
CDC	$8(\log_2(N_{\text{FFT}})+1)$

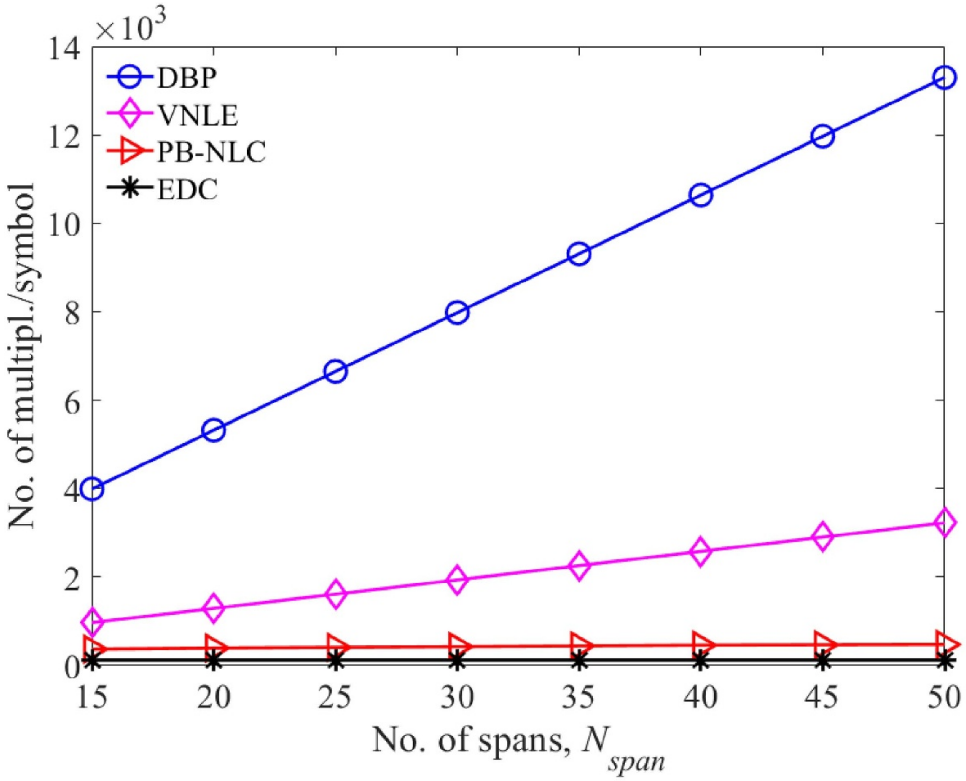

Figure 11. The number of real-valued multiplications/symbol for DBP, VNLE, PB-NLC, and EDC techniques as a function of the number of fiber spans N_{span} for a quasi-Nyquist single-carrier system.

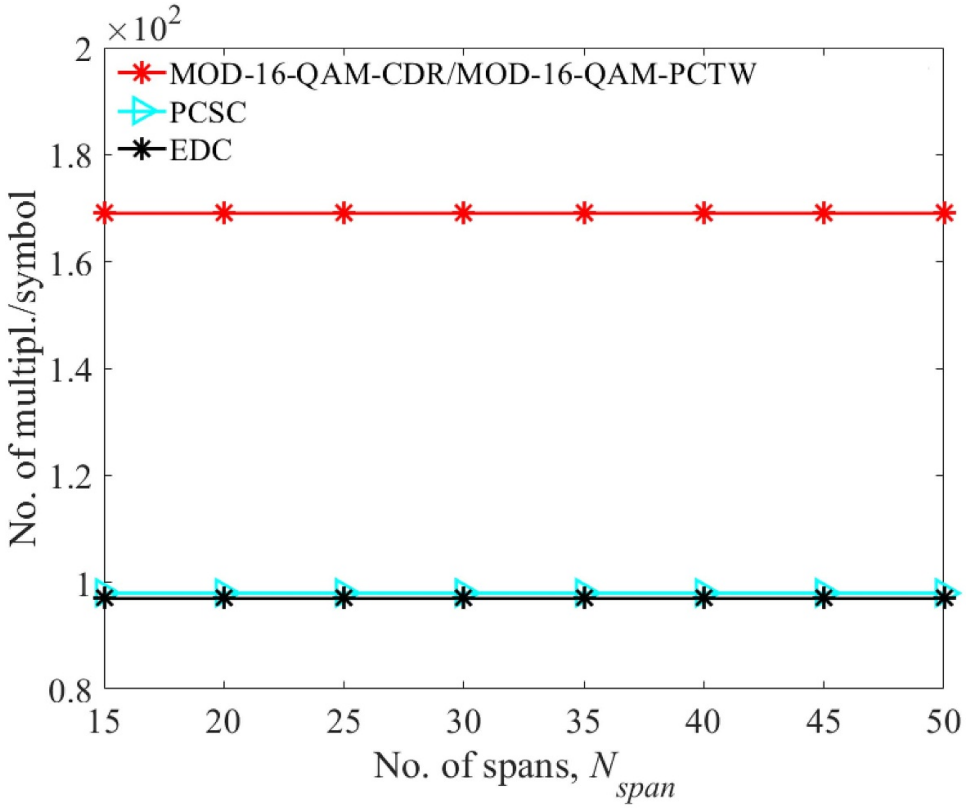

Figure 12. The number of real-valued multiplications/symbol for MOD-16-QAM-CDR/MOD-16-QAM-PCTW, PCSC, and EDC techniques as a function of the number of fiber spans N_{span} for a CO-OFDM system.

DBP has been carried out. The results indicate that the non-deterministic distributed PMD impairs the DBP performance larger than the Kerr-induced signal-ASE stochastic nonlinear effect and manifests as a fundamental limitation to the achievable fiber channel capacity. In [74], the impact of PMD on the performance of the in-line phase-conjugation technique has been studied. The study results indicate that a more close installation of the in-line phase-conjugation devices significantly improves the NLC performance. Similar to PMD, several research works have been conducted in the literature to investigate the impact of laser phase noise on NLC performance. In [75], analytical modeling of the equalization enhanced phase noise (EEPN) has been developed for evaluating the performance of the optical transmission systems. The findings in [75] demonstrate the importance of considering EEPN in the design of the long-haul optical transmission system. In [76], an experimental study on the impact of EEPN on the optical transmission system has been carried out. The results indicate that the blind phase search algorithm used to recover the carrier phase can partially mitigate the effect of EEPN in an optical transmission system.

9. Conclusion

In this tutorial, we have elucidated the impact of Kerr-induced fiber nonlinearity effects on long-haul coherent optical transmission systems. Using the classical electron oscillator model, we have explained the origin of the nonlinear susceptibility and the Kerr effect in a silica-based optical fiber. We have discussed the mechanism of the optical pulse propagation in the optical fiber medium using the nonlinear Schrödinger equation derived from Maxwell's equations. Then, we have investigated the numerical and analytical methods commonly used to solve the pulse propagation equation, such as the SSFM, the Volterra series-based analysis, and the regular perturbation series-based analysis. Following that, we have demonstrated various state-of-the-art fiber NLC techniques available in the literature to deal with the detrimental effects of fiber Kerr nonlinearity. We have also carried out the performance comparison and the complexity evaluation of the selected NLC techniques for quasi-Nyquist single-carrier and CO-OFDM transmission systems.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

ORCID iD

Sunish Kumar Orappanpara Soman
https://orcid.org/0000-0002-1010-364X

References

[1] Agrawal G P 1995 *Nonlinear Fiber Optics* (San Diego: Academic)
[2] Agrawal G P 2000 *Nonlinear Fiber Optics* (Berlin: Springer)
[3] Soma D *et al* 2018 *J. Lightwave Technol.* **36** 1362
[4] Chen M, Saad W Yin C and Debbah M 2019 *IEEE Trans. Commun.* **67** 4267
[5] Zhang J, Rajendran S, Sun Z, Woods R and Hanzo L 2019 *IEEE Wirel. Commun.* **26** 92
[6] Restuccia F, D'Oro S, Melodia T and Hanzo L 2018 *IEEE Internet Things J.* **5** 4829
[7] Amari A *et al* 2017 *IEEE Commun. Surv. Tutorials* **19** 3097
[8] Xu C, Liu X and Wei X 2004 *IEEE J. Sel. Top. Quantum Electron.* **10** 281
[9] Laperle C, Villeneuve B, Zhang Z and McGhan D 2008 *J. Lightwave Technol.* **26** 168
[10] Sunish Kumar O S, Amari A, Dobre O A and Venkatesan R 2018 *IEEE Photonics Technol. Lett.* **30** 769
[11] Amari A *et al* 2020 *IEEE Photonics Technol. Lett.* **32** 1435
[12] Nguyen T T, Harper P, Sunish Kumar O S and Ellis A 2020 *Proc. Opt. Fiber Commun. Conf. Exhib.* (San Diego: OSA)
[13] Kumar O S S *et al* 2017 *Proc. Asia Commun. Photonics Conf.* (Guangzhou: IEEE)
[14] Rios-Müller R *et al* 2015 *J. Lightwave Technol.* **33** 1402
[15] Rios-Müller R *et al* 2015 *Proc. Eur. Conf. Opt. Commun.* (Spain: IEEE)
[16] Chandrasekhar S and Liu X 2012 *J. Lightwave Technol.* **30** 3816
[17] Xia Z, Cui S, Ke C, Fu S, Tang M and Liu D 2016 *Proc. Opt. Fiber Commun. Conf. Exhib.* (Anaheim: OSA)
[18] Killey R I, Watts P M, Mikhailov V, Glick M and Bayvel P 2005 *IEEE Photonics Technol. Lett.* **17** 714
[19] Kuschnerov M *et al* 2009 *J. Lightwave Technol.* **27** 3614
[20] Renaudier J *et al* 2010 *Bell Lab. Tech. J.* **14** 27
[21] Savory S J 2006 *Bell Lab. Tech. J.* **42** 407
[22] Kumar O S S *et al* 2014 *Proc. Int. Conf. Signal Process. Integr. Netw.* (Noida: IEEE)
[23] Meena D *et al* 2013 *Proc. Intern. Conf. on Fiber Opti. and Photon.* (Chennai: OSA)
[24] Rafique D 2016 *J. Lightwave Technol.* **34** 544
[25] Ding M *et al* 2018 *Basics of Optical Fiber Measurements* (Singapore: Springer)
[26] Wilder M R 2000 *Fiberoptics Online* (available at: www.fiberopticsonline.com/doc/understanding-and-measuring-chromatic-dispers-0002)
[27] Kumar S and Jamal M 2014 *Fiber Optic Communications: Fundamentals and Applications* (New Jersey: Wiley)
[28] Weiss G H and Maradudin A A 1962 *J. Math. Phys.* **3** 771
[29] Guiomar F P *et al* 2011 *IEEE Photonics Technol. Lett.* **23** 1412
[30] Amari A *et al* 2017 *Proc. Int. Conf. Transparent Opt. Netw.* (Spain: IEEE)
[31] Vgenopoulou V *et al* 2014 *Proc. Asia Commun. Photonics Conf.* (Shanghai: IEEE)
[32] Bakhshali A *et al* 2016 *J. Lightwave Technol.* **34** 1770
[33] Schetzen M 1980 *The Volterra and Wiener Theories of Nonlinear Systems* (New Jersey: Wiley)
[34] Peddanarappagari K V and Brandt-Pearce M 1997 *J. Lightwave Technol.* **15** 2232
[35] Diamantopoulos N P *et al* 2019 *J. Lightwave Technol.* **37** 1214
[36] Vannucci A, Serena P and Bononi A 2002 *J. Lightwave Technol.* **20** 1102
[37] Mecozzi A, Clausen C B and Shtaif M 2000 *IEEE Photonics Technol. Lett.* **12** 392
[38] Tao Z *et al* 2011 *J. Lightwave Technol.* **29** 2570
[39] Gao Y *et al* 2014 *Opt. Express* **22** 1209
[40] Kumar O S S *et al* 2019 *IEEE Photonics J.* **11** 1
[41] Ellis A D *et al* 2017 *Adv. Opt. Photonics* **9** 429
[42] Xu T *et al* 2021 *Opt. Express* **29** 17428
[43] Vassilieva O *et al* 2019 *J. Lightwave Technol.* **37** 50
[44] Gordon J P and Mollenauer L F 1990 *Opt. Lett.* **15** 1351
[45] Ghazisaeidi A 2017 *J. Lightwave Technol.* **35** 5150

[46] Kumar O S S *et al* 2020 Intra-channel nonlinearity compensation based on second-order perturbation theory (arXiv:2005.01191v1 [cs.IT])
[47] Kumar O S S *et al* 2021 Deep neural network assisted second-order perturbation-based nonlinearity compensation (arXiv:2105.09027 [eess.SP])
[48] Orappanpara Soman Sunish Kumar 2020 Digital signal processing techniques for fiber nonlinearity compensation in coherent optical communication systems *PhD Thesis* Memorial University of Newfoundland, 230 Elizabeth Ave, St. John's, NL A1C 5S7, Canada (https://doi.org/10.48336/e70n-7d17)
[49] Soman S K O *et al* 2021 *J. Lightwave Technol.* **39** 5474
[50] Liu X *et al* 2013 *Nat. Photon.* **7** 560
[51] Liu X *et al* 2014 *Proc. Opto-Electron. Commun. Conf.* (Melbourne: OSA)
[52] Eliasson H *et al* 2015 *Opt. Express* **23** 2392
[53] Yi X *et al* 2015 *Opt. Express* **22** 13454
[54] Le S T *et al* 2015 *J. Lightwave Technol.* **33** 1308
[55] Yoshida T *et al* 2014 *Proc. Opt. Fiber Commun. Conf. Exhib.* (San Francisco: OSA)
[56] Le S T *et al* 2015 *J. Lightwave Technol.* **33** 2206
[57] Kumar O S S *et al* 2017 *Proc. SPIE* **10130** 1–10
[58] Foursa D G *et al* 2013 *Proc. Opt. Fiber Commun. Conf. Exhib.* (Anaheim: OSA)
[59] Jansen S L *et al* 2006 *J. Lightwave Technol.* **24** 54
[60] Ellis A and Sorokina M 2019 *Optical Communication Systems: Limits and Possibilities* (Singapore: Jenny Stanford Publishing)
[61] Liu X *et al* 2013 *Proc. Opt. Fiber Commun. Conf. Exhib.* (Anaheim: OSA)
[62] Sinkin O V *et al* 2003 *J. Lightwave Technol.* **21** 61
[63] Rafique D *et al* 2011 *Opt. Express* **19** 9453
[64] Li X *et al* 2008 *Opt. Express* **16** 880
[65] Rafique D *et al* 2011 *Proc. Int. Conf. Transparent Opt. Netw.* (Sweden: IEEE)
[66] Li L *et al* 2011 *Proc. Opt. Fiber Commun. Conf. Exhib.* (Los Angeles: OSA)
[67] Irukulapati N V *et al* 2014 *IEEE Trans. Commun.* **62** 3956
[68] Liga G *et al* 2014 *Opt. Express* **22** 30053
[69] Fontaine N K 2013 *Proc. Eur. Conf. Opt. Commun.* (London: IEEE)
[70] Maher R *et al* 2015 *Proc. Opt. Fiber Commun. Conf. Exhib.* (Los Angeles: OSA)
[71] Guiomar F P *et al* 2016 *J. Lightwave Technol.* **34** 1896
[72] Zhang F *et al* 2015 *J. Lightwave Technol.* **33** 5140
[73] Gao G *et al* 2012 *Opt. Express* **20** 14406
[74] McCarthy M E *et al* 2016 *Opt. Express* **24** 3385
[75] Jin C *et al* 2021 *J. Lightwave Technol.* **39** 4646
[76] Arnould A and Ghazisaeidi A 2019 *J. Lightwave Technol.* **37** 5282